**PCT** WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07K 14/705 — A1

(11) International Publication Number: **WO 99/21890**

(43) International Publication Date: 6 May 1999 (06.05.99)

(21) International Application Number: PCT/SE98/01947

(22) International Filing Date: 27 October 1998 (27.10.98)

(30) Priority Data:

| | | |
|---|---|---|
| 9703914-3 | 27 October 1997 (27.10.97) | SE |
| 9800864-2 | 16 March 1998 (16.03.98) | SE |
| 9802575-2 | 17 July 1998 (17.07.98) | SE |

(71) Applicant *(for all designated States except US)*: ASTRA AKTIEBOLAG [SE/SE]; S–151 85 Södertälje (SE).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: EKSTRAND, Jonas [SE/SE]; Astra Hässle AB, Tvistevägen 48, S–907 36 Umeå (SE).

(74) Agent: ASTRA AKTIEBOLAG; Patent Dept., S–151 85 Södertälje (SE).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published**

*With international search report.*

*Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.*

(54) Title: NEW NUCLEOTIDE SEQUENCES

(57) Abstract

The present invention relates to a nucleic acid molecule encoding a $GABA_B$ receptor, or a functionally equivalent modified form thereof, said receptor being selected from the group consisting of human and canine $GABA_B$ receptors.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **AL** | Albania | **ES** | Spain | **LS** | Lesotho | **SI** | Slovenia |
| **AM** | Armenia | **FI** | Finland | **LT** | Lithuania | **SK** | Slovakia |
| **AT** | Austria | **FR** | France | **LU** | Luxembourg | **SN** | Senegal |
| **AU** | Australia | **GA** | Gabon | **LV** | Latvia | **SZ** | Swaziland |
| **AZ** | Azerbaijan | **GB** | United Kingdom | **MC** | Monaco | **TD** | Chad |
| **BA** | Bosnia and Herzegovina | **GE** | Georgia | **MD** | Republic of Moldova | **TG** | Togo |
| **BB** | Barbados | **GH** | Ghana | **MG** | Madagascar | **TJ** | Tajikistan |
| **BE** | Belgium | **GN** | Guinea | **MK** | The former Yugoslav Republic of Macedonia | **TM** | Turkmenistan |
| **BF** | Burkina Faso | **GR** | Greece | | | **TR** | Turkey |
| **BG** | Bulgaria | **HU** | Hungary | **ML** | Mali | **TT** | Trinidad and Tobago |
| **BJ** | Benin | **IE** | Ireland | **MN** | Mongolia | **UA** | Ukraine |
| **BR** | Brazil | **IL** | Israel | **MR** | Mauritania | **UG** | Uganda |
| **BY** | Belarus | **IS** | Iceland | **MW** | Malawi | **US** | United States of America |
| **CA** | Canada | **IT** | Italy | **MX** | Mexico | **UZ** | Uzbekistan |
| **CF** | Central African Republic | **JP** | Japan | **NE** | Niger | **VN** | Viet Nam |
| **CG** | Congo | **KE** | Kenya | **NL** | Netherlands | **YU** | Yugoslavia |
| **CH** | Switzerland | **KG** | Kyrgyzstan | **NO** | Norway | **ZW** | Zimbabwe |
| **CI** | Côte d'Ivoire | **KP** | Democratic People's Republic of Korea | **NZ** | New Zealand | | |
| **CM** | Cameroon | | | **PL** | Poland | | |
| **CN** | China | **KR** | Republic of Korea | **PT** | Portugal | | |
| **CU** | Cuba | **KZ** | Kazakstan | **RO** | Romania | | |
| **CZ** | Czech Republic | **LC** | Saint Lucia | **RU** | Russian Federation | | |
| **DE** | Germany | **LI** | Liechtenstein | **SD** | Sudan | | |
| **DK** | Denmark | **LK** | Sri Lanka | **SE** | Sweden | | |
| **EE** | Estonia | **LR** | Liberia | **SG** | Singapore | | |

# NEW NUCLEOTIDE SEQUENCES

## TECHNICAL FIELD

The present invention relates in particular to nucleic acid molecules encoding $GABA_B$ receptors, and to methods for screening for compounds which are inhibitors of transient lower esophageal sphincter relaxations (TLESR), said methods comprising the use of a nucleic acid molecule encoding a $GABA_B$ receptor.

## BACKGROUND ART

### *$GABA_B$ receptors*

GABA (4-aminobutanoic acid) is an endogenous neurotransmitter in the central and peripheral nervous systems. Receptors for GABA have traditionally been divided into $GABA_A$ and $GABA_B$ receptor subtypes. $GABA_B$ receptors (for a review see Kerr, D.I.B. and Ong, J. (1995) Pharmac. Ther. vol. 67, pp.187-246) belong to the superfamily of G-protein coupled receptors. $GABA_B$ receptor agonists are described as being of use in the treatment of CNS disorders, such as muscle relaxation in spinal spasticity, cardiovascular disorders, asthma, gut motility disorders such as irritable bowel syndrome and as prokinetic and anti-tussive agents. $GABA_B$ receptor agonists have also been disclosed as useful in the treatment of emesis (WO 96/11680).

The cloning of the rat $GABA_B$ receptors $GABA_BR1a$ (SEQ ID NOS: 44 and 45) and $GABA_BR1b$ (SEQ ID NOS: 46 and 47) is disclosed by Kaupmann et al. (1997) Nature, vol. 386, 239-246. The mature rat $GABA_BR1b$ differed from $GABA_BR1a$ in that the N-terminal 147 residues were replaced by 18 different residues. It was presumed that the rat $GABA_BR1a$ and -b receptor variants are derived from the same gene by alternative splicing.

The cloning of the human $GABA_B$ receptor $GABA_BR1b$ is disclosed in WO 97/46675.

*Reflux*

In some humans, the lower esophageal sphincter (LES) is prone to relaxing more frequently than in other humans. As a consequence, fluid from the stomach can pass into the esophagus since the mechanical barrier is temporarily lost at such times, an event hereinafter referred to as "reflux".

Gastro-esophageal reflux disease (GERD) is the most prevalent upper gastrointestinal tract disease. Current therapy has aimed at reducing gastric acid secretion, or by reducing esophageal acid exposure by enhancing esophageal clearance, lower esophageal sphincter tone and gastric emptying. The major mechanism behind reflux has been considered to depend on a hypotonic lower esophageal sphincter. However, recent research (e.g. Holloway & Dent (1990) Gastroenterol. Clin. N. Amer. 19, 517-535) has shown that most reflux episodes occur during transient lower esophageal sphincter relaxations (TLESR), i.e. relaxations not triggered by swallows. It has also been shown that gastric acid secretion usually is normal in patients with GERD. Consequently, there is a need for compounds which reduce the incidence of TLESR and thereby prevent reflux.

## DISCLOSURE OF THE INVENTION

In the applicants' earlier patent application WO 98/11885, filed on 15 September 1997, it is disclosed that $GABA_B$ receptor agonists can be used to reduce the incidence of transient lower esophageal sphincter relaxations (TLESR).

The present invention provides nucleic acid molecules encoding human and canine $GABA_B$ receptors. These nucleic acid molecules will make possible the screening for compounds which are agonists or antagonists of $GABA_B$ receptors, e.g. compounds which are inhibitors of transient lower esophageal sphincter relaxations (TLESR).

Consequently, in a first aspect, the present invention provides a nucleic acid molecule encoding a $GABA_B$ receptor, or a functionally equivalent modified form thereof, said receptor being selected from the group consisting of human and canine $GABA_B$ receptors.

In preferred forms of the invention, the said nucleic acid molecule encodes the human $GABA_B$ receptor 1a (SEQ ID NOS: 48 and 49), 1b (SEQ ID NOS: 50 and 51), 1c (SEQ ID NOS: 54 and 55) or 1d (SEQ ID NOS: 56 and 57); or the canine $GABA_B$ receptor 1a (SEQ ID NOS: 52 and 53) or 1c (SEQ ID NOS: 58 and 59). Accordingly, the invention furthermore provides a nucleic acid molecule selected from:

(a) nucleic acid molecules comprising a nucleotide sequence set forth as SEQ ID NO: 48, 50, 52, 54, 56 or 58;

(b) nucleic acid molecules comprising a nucleotide sequence capable of hybridizing, under stringent hybridization conditions, to a nucleotide sequence complementary to the polypeptide coding region of a DNA molecule as defined in (a); and

(c) nucleic acid molecules comprising a nucleotide sequence which is degenerate as a result of the genetic code to a nucleotide sequence as defined in (a) or (b).

Furthermore, the invention provides nucleic acid molecules of genomic origin encoding human $GABA_B$ receptors (SEQ ID NOS: 60 and 61) as well as nucleic acid molecules (set forth as SEQ ID NO: 70, 72, 74, 76, 78, 80, 82, 84) encoding additional isoforms of the human $GABA_B$ receptor which isoforms can be shown to be generated by alternative splicing.

It should thus be understood that the nucleic acid molecule according to the invention is not to be limited strictly to molecules comprising the sequences set forth as SEQ ID : 48, 50, 52, 54, 56 or 58. Rather the invention encompasses nucleic acid molecules carrying modifications like substitutions, small deletions, insertions or inversions, which nevertheless encode proteins having substantially the biochemical activity of the $GABA_B$ receptors according to the invention. Included in the invention are consequently nucleic acid molecules, the nucleotide sequence of which is at least 95% homologous, preferably at least 96%, 97%, 98% or 99% homologous, with the nucleotide sequence shown as SEQ ID NO: 48, 50, 52, 54, 56 or 58 in the Sequence Listing.

The term "stringent hybridization conditions" is known in the art from standard protocols (e.g. Current Protocols in Molecular Biology, editors F. Ausubel et al., John Wiley and Sons, Inc. 1994) and could be understood as as stringent or more stringent than those defined by e.g. hybridization to filter-bound DNA in 0.5 M $NaHPO_4$, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at +65°C, and washing in 0.1xSSC / 0.1% SDS at +68°C.

Included in the invention is also a nucleic acid molecule which nucleotide sequence is degenerate, because of the genetic code, to a nucleic acid of the present invention and more particularly to one of the nucleotide sequences set forth as SEQ ID NOs: 48, 50, 52, 54, 56 and 58. A sequential grouping of three nucleotides, a "codon", codes for one amino acid. Since there are 64 possible codons, but only 20 natural amino acids, most amino acids are coded for by more than one codon. This natural "degeneracy", or "redundancy", of the genetic code is well known in the art. It will thus be appreciated that the nucleic acid sequences shown in the Sequence Listing is only an example within a large but definite group of nucleic acid sequences which will encode the polypeptide as described above.

In a further aspect, the invention provides a recombinant polypeptide encoded by a nucleotide sequence of the present invention, encoding a $GABA_B$ receptor. In preferred forms of the inventions, the said polypeptide comprises an amino acid sequence set forth as SEQ ID NO: 49, 51, 53, 55, 57, 59, 71, 73, 75, 77, 79, 81, 83 or 85 in the Sequence Listing. However, polypeptides of the present invention are not to limited to those having an amino acid sequence identical with one of SEQ ID NOs: 49, 51, 53, 55, 59, 71, 73, 75, 77, 79, 81, 83 or 85 in the Sequence Listing. Rather the invention encompasses polypeptides carrying modifications such as substitutions, small deletions, insertions or inversions, which polypeptides nevertheless have substantially the biological activities of the $GABA_B$ receptor. Included in the invention are consequently polypeptides, the amino acid sequence of which is at least 95% homologous, preferably at least 96%, 97%, 98% or 99% homologous, with one of the amino acid sequences described by SEQ ID NOs: 49, 51, 53, 55, 57 59, 71, 73, 75, 77, 79, 81, 83 and 85 in the Sequence Listing.

Included in the invention are polypeptides of the present invention which have been post-translationally modified, e.g. by cleavage of an N-terminal signal sequence which can be e.g. 1 to 25 amino acids long.

In yet another aspect, the invention provides a vector transformed with a nucleic acid molecule of the present invention. The said vector can e.g. be a replicable expression vector which carries and is capable of mediating the expression of a nucleic acid molecule according to the invention. In the present context the term "replicable" means that the vector is able to replicate in a given type of host cell into which is has been introduced. Examples of vectors are viruses such as bacteriophages, cosmids, plasmids and other recombination vectors. Nucleic acid molecules are inserted into vector genomes by methods well known in the art.

Included in the invention is also a cultured host cell harbouring a vector according to the invention. Such a host cell can be a prokaryotic cell, a unicellular eukaryotic cell or a cell derived from a multicellular organism. The host cell can thus e.g. be a bacterial cell such as an *E. coli* cell; a cell from a yeast such as *Saccharomyces cervisiae* or *Pichia pastoris*, or a mammalian cell. The methods employed to effect introduction of the vector into the host cell are standard methods well-known to a person familiar with recombinant DNA methods.

A further aspect of the invention is a process for production of a $GABA_B$ receptor polypeptide according to the invention, said process comprising culturing a host cell as defined above under conditions whereby the said polypeptide is produced, and recovering the said polypeptide.

A further important aspect of the invention is a method for the screening of compounds which are inhibitors of transient lower esophageal sphincter relaxations (TLESR), said method comprising the use of a nucleic acid molecule encoding a $GABA_B$ receptor. The said nucleic acid molecule encoding a $GABA_B$ receptor can e.g. be one of the nucleic acid molecules according to the invention encoding human or canine $GABA_B$ receptors. However, it should be understood that this aspect of the invention is not limited to the use

of the said human and canine $GABA_B$ receptors, but rather encompasses the use of any $GABA_B$ receptor for screening for compounds which are inhibitors of transient lower esophageal sphincter relaxations.

In yet another important aspect, the invention provides a method for the screening of compounds which are agonists or antagonists to a $GABA_B$ receptor, said method comprising the use of a nucleic acid molecule, according to the invention, encoding human or canine $GABA_B$ receptors.

**Brief Description of the Drawings**

Figure 1: Map of the human $GABA_B$ receptor gene.

The exon/intron organisation is shown. Exons are indicated as solid boxes numbered 1-23. The part of intron 5 that is retained together with exon 6 giving rise to $GABA_B$ receptor 1b is indicated as an open box.

Figure 2: Expression of human $GABA_B$ receptor 1b isoform in transfected C127 cells.

Western blot analysis of transfected C127 cells using a polyclonal anti-human $GABA_B$ receptor antibody. Lane 1: Untransfected C127 whole cell lysate. Lanes 2-7: Whole cell lysates of six independent clones transfected with human $GABA_B$ receptor 1b isoform encoding cDNA. The clones analysed in lanes 4 to 7 express a $GABA_B$ receptor of expected molecular weight (arrow).

Figure 3: Expression of human $GABA_B$ receptor 1d isoform in transfected C127 cells.

Western blot analysis of transfected C127 cells using a polyclonal anti-human $GABA_B$ receptor antibody. Lanes 1-3: Concentrated culture media from three independent C127 clones transfected with a cDNA expression construct encoding the human $GABA_B$ receptor

1d isoform. Lanes 4-6: Whole cell lysates corresponding to the clones analysed in lanes 1-3. The experiment revealed that the human $GABA_B$ receptor 1d cDNA encodes a secreted isoform. The arrow indicates the bands corresponding to the 1d isoform.

Figure 4: Expression of human $GABA_B$ receptor 1d isoform in *E. coli*.

Western blot analysis of transformed *E. coli* cells using a polyclonal anti-human $GABA_B$ receptor antibody. Lane 1: Lysate from an uninduced *E. coli* culture transformed with an pET-based expression construct encoding the human $GABA_B$ receptor 1d cDNA. Lane 2: Lysate from an IPTG-induced *E. coli* culture transformed with an expression construct encoding the human $GABA_B$ receptor 1d cDNA. Lane 3: Lysate from an IPTG-induced *E. coli* culture transformed with an expression construct encoding an unrelated protein. Lane 4: The BSA-conjugated peptide previously used for immunization was loaded on the gel as a positive antibody control.

The screening methods according to the invention can e.g. comprise the steps (a) transforming a cultured cell with a nucleic acid molecule encoding a $GABA_B$ receptor, so that a $GABA_B$ receptor is expressed on the surface of the cell; (b) contacting a test compound with the said cell; and (c) determining whether the test compound binds to, and/or activates, the $GABA_B$ receptor.

In particular $GABA_B$ receptor expressing cells, transgenic animals or cells and tissues derived thereof, nay be used to screen substance libraries for antagonist or agonist activities. For this purpose, $GABA_B$ receptor expression may be directed to cells and tissues containing, either naturally or artificially, the necessary components allowing correct receptor transport and processing as well as coupling to second messenger pathways. Screening may be performed as ligand binding assays or functional assays. For screening, cells and tissues my be prepared in various ways, each uniquely suited to its purpose. Ligand binding assays are performed *in vivo* or *in vitro* using e.g. radiolabelled GABA. Functional assays examplified by, but not limited to, $Ca^{++}$-responses, cAMP-

responses and effects on $K^+$ channels, may be performed in living cells, broken cells, isolated cell membranes, as well as in tissues and in living animals. To facilitate measurement of physiological $GABA_B$ receptor mediated responses, $GABA_B$ receptors may be co-expressed with promiscous G-proteins like e.g. Gα16 or Gqi5 increasing G-protein coupling. Another possible way to increase G-protein coupling is to fuse the $GABA_B$ receptor with appropriate G-proteins using standard molecular techniques. In order to further improve readouts in $Ca^{++}$-response assays, $GABA_B$ receptors may be co-expressed with aequorin, a photoprotein cloned from the luminescent jellyfish *Aequorea victoria*.

In yet another aspect, the invention provides a pharmaceutical compositions comprising a soluble $GABA_B$ receptor further comprising one or more pharmaceutical acceptable carriers and/or diluents.

In yet another aspect, the invention provides a method for the diagnosis or treatment of conditions related to GABA-dysfunction, e.g. epilepsy, psychiatric disorders such as depression and anxiety, cognitive dysfunction, gastroesophageal reflux disease, emesis, irritable bowel syndrome, dyspepsia, spasticity, arthritis, allergies, autoimmune diseases, neoplastic diseases, pain, and infectious diseases, comprising the use or administration of soluble forms of the $GABA_B$ receptor, such as the human $GABA_B$ receptor 1c or 1d or functionally equivalent modified forms thereof; preferably being a polypeptidfe of the present invention..

The soluble forms of the $GABA_B$ receptor can e.g. be produced by culturing a host cell harbouring a vector comprising a nucleic acid encoding the soluble $GABA_B$ receptor under conditions where whereby the said polypeptide is produced, the polypeptide recovered and administered to a patient in need thereof.

Furthermore, the level of soluble $GABA_B$ receptors in certain body fluids, e.g. serum and CSF, can be measured and used in diagnostics related to conditions with altered levels of soluble $GABA_B$ receptors in said body fluids.

Throughout this description the terms "standard protocols" and "standard procedures", when used in the context of molecular cloning techniques, are to be understood as protocols and procedures found in an ordinary laboratory manual such as: Current Protocols in Molecular Biology, editors F. Ausubel et al., John Wiley and Sons, Inc. 1994, or Sambrook, J., Fritsch, E.F. and Maniatis, T., Molecular Cloning: A laboratory manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY 1989.

EXAMPLES

EXAMPLE 1A: Cloning and sequencing of cDNA encoding human $GABA_B$ receptor 1a and 1b

Messenger RNA from human hippocampus was obtained from Clontech (Palo Alto, CA, USA) (catalog #6578-1). First-strand cDNA synthesis reactions were performed using the First-strand cDNA Synthesis kit from Amersham Pharmacia Biotech (Uppsala, Sweden). The $pd(N)_6$ primer was used to prime the first-strand synthesis. The generated cDNA molecules were used as template in the PCR reactions described below.

Specific PCR primers were designed (Table 1) based on the sequences of the rat $GABA_B$ receptor 1a and 1b cDNA (Kaupmann et al. 1997, EMBL accession numbers Y10369 (SEQ ID NO: 44) and Y10370 (SEQ ID NO: 46)).

TABLE 1

Primers used for RT-PCR on mRNA from human hippocampus

| Nr. | Species | Sequence 5´- 3´ | SEQ ID NO |
|---|---|---|---|
| 794 | Rat | GTTTCTTCTCGGATCCAGCTGTGCCTG | 1 |
| 795 | Rat | CAGGCACAGCTGGATCCGAGAAGAAACT | 2 |
| 796 | Rat | CGGTCGACTCACTTGTAAAGCAAATGTACTCGACTCCC | 3 |
| 797 | Rat | ATGCGCGCCGGCAGCCAACATGCTGCTGCTGCTGCTGGTGC | 4 |
| 831 | Rat | CGGTCGACTCACTTGTAAAGCAAATGTACTCGACTCCCATCACAGC | 5 |
| 838 | Rat | ATGCGCGCCGGCAGCCAACATGCTGCTGCTGCTGCTGGTGCCTCTCTTCC | 6 |
| 842 | Rat | CAGGCACAGCTGGATCCGAGAAGAAACTCTGTCGGAAAGT | 7 |
| 863 | Rat | GGTCATCCAGCGTTGAGGTGAAGAC | 8 |
| 864 | Rat | GAAGGTTGCCAGATTATACATCCGC | 9 |
| 865 | Rat | CCACGATGATTCGAGCATCTTGACG | 10 |
| 866 | Rat | GCCTCTCACTCCCCTCATCTCC | 11 |
| 932 | Human | GAGTGAAGGAGGCTGGAATTG | 12 |

cDNA fragments encoding parts of the human $GABA_B$ receptors were amplified directly by PCR using the designed primers with the generated cDNA molecules as templates. All PCR experiments were carried out using the Perkin Elmer Taq DNA polymerase with Gene Amp (Roche Molecular Systems Inc., NJ, USA) with the following PCR program: +95°C for 1 minute, +50°C for 30 seconds, +72°C for 3 minutes, repeated 44 times and finally +72°C for 7 minutes. The following primer combinations gave PCR products with expected sizes: 838 and 842, 838 and 795, 797 and 865, 864 and 865, 864 and 863 which corresponds to the 5´-end of the $GABA_B$ receptor 1a. Primer pair 932 and 831, 932 and 796, 794 and 831 gave PCR products which correspond to the 3´-end of both $GABA_B$ receptor 1a and 1b. The primer combination 839 and 918 gave a PCR product corresponding to the 5'-end of the $GABA_B$ receptor 1b.

The PCR products were subcloned into the pGEM-T vector from Promega (Madison, USA). The inserts were subjected to nucleotide sequence analysis, the complete nucleotide sequence for all subclones were determined using a Thermo Sequenase dye terminator cycle sequencing pre-mix kit (Amersham Pharmacia Biotech, Uppsala, Sweden). As primers for sequencing reactions specific oligonucleotides complementary to the vector pGEM-T or primers complementary to the cDNA encoding the $GABA_B$ receptor were used.

Additional PCR primers were designed based on the obtained sequences encoding fragments of the human $GABA_B$ receptor and additional DNA fragments encoding parts of the human $GABA_B$ receptors were amplified by PCR, the PCR products subcloned and sequenced as described above.

EXAMPLE 1B : Cloning and sequencing of the 3'-ends of the cDNA encoding human $GABA_B$ receptors 1a and 1b

Messenger RNA from human hippocampus was obtained from Clontech (Palo Alto, CA, USA) (catalogue #6578-1). First strand cDNA synthesis reactions were performed using the First-strand cDNA Synthesis kit from Amersham Pharmacia Biotech (Uppsala, Sweden). The Not I-$d(T)_{18}$ primer was used to prime the first-strand synthesis. The generated cDNA molecules were used as template in the PCR reactions described below.

Specific PCR primers were designed (Table 2 ) based on the sequences of the human $GABA_B$ receptor 1a and 1b cDNA obtained in Example 1A and the EST sequence EMBL accession number Y11044.

By homology searches in the EMBL database using the $GABA_B$ receptor cDNA sequences obtained in Example 1A as query the EST sequence EMBL accession number Y11044 have been found to be homologous to the 3'-end of the $GABA_B$ receptor cDNA.

TABLE 2

Primers used in PCR to amplify 3'ends of human $GABA_B$ receptor cDNA

| Nr | Species | Sequence 5´- 3´ | SEQ ID NO |
|---|---|---|---|
| 938 | Human | GACGCTTATCGAGCAGCTTC | 13 |
| 972 | Human | AGCCCAGAACTCACAGGGGGACAT | 14 |
| 973 | Human | GCTTCAAGCCAGGTACGAACTAA | 15 |

cDNA fragments encoding parts of the human $GABA_B$ receptors were amplified directly by PCR using the designed primers with the generated cDNA molecules as templates. All

PCR experiments were carried out using the Perkin Elmer Taq DNA polymerase with Gene Amp (Roche Molecular Systems Inc., NJ, USA) with the following PCR program: + 95°C for 1 minute, +50°C for 30 seconds,+72°C for 3 minutes, repeated 44 times and finally +72°C for 7 minutes. The following primer combinations gave PCR products with expected sizes: 938 and 972, 938 and 973 corresponding to the 3´end of both $GABA_B$ receptor 1a and 1b cDNA.

The PCR products were subcloned into the pGEM-T vector from Promega (Madison, USA). The inserts were subjected to nucleotide sequence analysis, the complete nucleotide sequence for all subclones were determined using a Thermo Sequenase dye terminator cycle sequencing pre-mix kit (Amersham Pharmacia Biotech, Uppsala, Sweden). As primers for sequencing reactions specific oligonucleotides complementary to the vector pGEM-T or primers complementary to the cDNA encoding the $GABA_B$ receptor were used.

EXAMPLE 1C : Cloning and sequencing of the 5'-end of the cDNA encoding human $GABA_B$ receptor 1b

Messenger RNA from human hippocampus was obtained from Clontech (Palo Alto, CA, USA) (catalogue #6578-1). Marathon cDNA amplification Kit (Clontech) was used for performing 5´/3'- RACE (Rapid Amplification of cDNA Ends). Adaptor-ligated double stranded cDNA molecules were amplified according to the manufacturers description. The $pd(N)_6$ primer from the First-strand cDNA Synthesis kit from Amersham Pharmacia Biotech (Uppsala, Sweden) was used to produce the adaptor-ligated cDNA.

A specific PCR primer was designed (Table 3) based on the sequences of the human $GABA_B$ receptor 1b cDNA obtained in Example 1A.

TABLE 3

Primers used in PCR to amplify the 5'-ends of human $GABA_B$ receptor cDNA

| Nr. | Species | Sequence 5´-3´ | SEQ ID NO |
|---|---|---|---|
| 958 | Human | TGGCCCTCCACCGCCTCAGTCATCTCA | 16 |
| AP1 | Marathon kit | CCATCCTAATACGACTCACTATAGGGC | 17 |

cDNA fragments encoding part of the human $GABA_B$ receptors were amplified directly by PCR using the designed primers with the generated adaptor-ligated cDNA molecules as template. The PCR experiment was carried out using the Expand Long Template PCR System (Boehringer Mannheim GmbH, Germany) with the following PCR program: + 94°C for 1minute, + 94°C for 30 seconds, +60°C for 30 seconds, +68°C for 4 minutes, repeated 24 times. The primer combination AP1 and 958 gave a PCR product which corresponded to the 5´-end of the $GABA_B$ receptor 1b cDNA, including 190 base pairs upstream the initiation codon.

The PCR products were subcloned into the pGEM-T vector from Promega (Madison, USA). The inserts were subjected to nucleotide sequence analysis, the complete nucleotide sequence for all subclones were determined using a Thermo Sequenase dye terminator cycle sequencing pre-mix kit (Amersham Pharmacia Biotech, Uppsala, Sweden). As primers for sequencing reactions specific oligonucleotides complementary to the vector pGEM-T or primers complementary to the cDNA encoding the $GABA_B$ receptor were used.

EXAMPLE 1D : Cloning and sequencing of the 5'-end of the cDNA encoding human $GABA_B$ receptor 1a

Messenger RNA from human hippocampus was obtained from Clontech (Palo Alto, USA) (catalogue #6578-1). Marathon cDNA amplification Kit (Clontech) was used to obtain adaptor-ligated double stranded cDNA molecules according to the manufacturer's description. The $pd(N)_6$ primer from the First-strand cDNA Synthesis kit from Amersham Pharmacia Biotech (Uppsala, Sweden) was used to obtain the adaptor-ligated cDNA .

Specific PCR primers were designed (Table 4) based on the sequences of the human $GABA_B$ receptor 1a cDNA obtained in Example 1 and the rat $GABA_B$ receptor 1a cDNA disclosed in WO 97/46675.

TABLE 4

Primers used to amplify 5'-ends of the human GABAB receptor 1a cDNA

| Nr. | Species | Sequence 5´-3´ | SEQ ID NO |
|---|---|---|---|
| 1033 | Human | CTCAATCTCATAGTCCACTGG | 18 |
| 1087 | Rat | CCTTGAGGCCCGGGGAGAG | 19 |

cDNA fragments encoding part of the human GABAB 1a receptor were amplified directly by PCR using the designed primers with the generated adaptor-ligated cDNA molecules as template. The PCR was performed using the Perkin Elmer Taq DNA polymerase with Gene Amp (Roche Molecular Systems Inc., NJ, USA) with the following PCR program: +94°C for 1 minute, +50°C for 30 seconds,+72°C for 3 minutes; +94°C for 1 minute, +60°C for 30 seconds,+72°C for 4 minutes, repeated 34 times and finally +72°C for 7 minutes.

The primer combination 1087 and 1033 gave a PCR product corresponding to the 5´-end of the $GABA_B$ receptor 1a cDNA , including 26 base pairs upstream the initiation codon.

The PCR products were subcloned into the pGEM-T vector from Promega (Madison, USA). The inserts were subjected to nucleotide sequence analysis, the complete nucleotide sequence for all subclones were determined using a Thermo Sequenase dye terminator cycle sequencing pre-mix kit (Amersham Pharmacia Biotech, Uppsala, Sweden). As primers for sequencing reactions specific oligonucleotides complementary to the vector pGEM-T were used.

A complete cDNA sequence encoding the human $GABA_B$ receptors 1a (SEQ ID NO: 48) and the human $GABA_B$ receptor 1b (SEQ ID NO: 50) were obtained by aligning the sequences of the different fragments cloned and sequenced in Examples 1A, 1B, 1C and 1D.

EXAMPLE 2A: Cloning and sequencing of cDNA encoding canine $GABA_B$ receptor 1a

QuickPrep Micro mRNA Purification kit (Amersham Pharmacia Biotech, Uppsala, Sweden) was used to isolate mRNA from canine neural tissues according to the manufacturers description. First-strand cDNA synthesis reactions were performed using the First-strand cDNA Synthesis kit from Amersham Pharmacia Biotech (Uppsala, Sweden). The Not-I-$d(T)_{18}$ bifunctional or $pd(N)_6$ primer was used to prime the first-strand synthesis. The generated cDNA molecules were used as template in the PCR reactions described below.

Specific PCR primers (Table 5) were designed based on the sequences of the rat $GABA_B$ receptor 1a and 1b cDNA (Kaupmann et al. 1997, EMBL accession numbers Y10369 (SEQ ID NO: 44 and Y10370 (SEQ ID NO: 46)).

TABLE 5

Primers used for RT-PCR on mRNA from canine cortex

| Nr. | Species | Sequence 5´-3´ | SEQ ID NO |
|---|---|---|---|
| 795 | Rat | CAGGCACAGCTGGATCCGAGAAGAAACT | 20 |
| 831 | Rat | CGGTCGACTCACTTGTAAAGCAAATGTACTCGACTCCCATCACAGC | 21 |
| 838 | Rat | ATGCGCGCCGGCAGCCAACATGCTGCTGCTGCTGCTGGTGCCTCTCTTCC | 22 |
| 840 | Rat | CGTCAAGATGCTCGAATCATCG | 23 |
| 841 | Rat | CAGGGGGCTCAGAGGGTCCC | 24 |
| 842 | Rat | CAGGCACAGCTGGATCCGAGAAGAAACTCTGTCGGAAAGT | 25 |
| 844 | Rat | CGGTCGACTCACTTGTAAAGCAAATGTACTCGACTCCCATCACAGCTAAG | 26 |
| 848 | Rat | ACTTTCCGACAGAGTTTCTTCTCGGATCCAGCTGTGCCTG | 27 |
| 865 | Rat | CCACGATGATTCGAGCATCTTGACG | 28 |

cDNA fragments encoding parts of the canine $GABA_B$ receptor were amplified directly by PCR using the designed primers with the generated cDNA molecules as templates. All PCR experiments were carried out using the Perkin Elmer Taq DNA polymerase with Gene Amp (Roche Molecular Systems Inc., NJ, USA) with the following PCR program: +95°C for 1 minute, +50°C for 30 seconds, +72°C for 3 minutes, repeated 44 times and finally

+72°C for 7 minutes. The following primer combinations gave PCR products with expected sizes: 842 and 838, 838 and 795, 838 and 865 which corresponds to the 5´-part of the canine $GABA_B$ receptor cDNA . Primer pair 848 and 844, 848 and 831, 848 and 841, 840 and 841 gave PCR products which correspond to the 3´-part of the canine $GABA_B$ receptor cDNA.

The PCR products were subcloned into the pGEM-T vector from Promega (Madison, USA). The inserts were subjected to nucleotide sequence analysis, the complete nucleotide sequence for all subclones were determined using a Thermo Sequenase dye terminator cycle sequencing pre-mix kit (Amersham Pharmacia Biotech, Uppsala, Sweden). As primers for sequencing reactions specific oligonucleotides complementary to the vector pGEM-T or primers complementary to the cDNA encoding the $GABA_B$ receptor were used.

EXAMPLE 2B: Cloning and sequencing of the 3'- and 5'-ends of the cDNA encoding canine $GABA_B$ receptor 1a

QuickPrep Micro mRNA Purification kit (Amersham Pharmacia Biotech, Uppsala, Sweden) was used to isolate mRNA from canine nerve tissues according to the manufactures description. Marathon cDNA amplification Kit (Clontech, Palo Alto, CA, USA) were used for performing both 5´-and 3´-RACE. Two adaptor-ligated double stranded cDNA libraries were amplified according to the manufacturers description. A random primer ($pd(N)_6$) was used when amplifying the adaptor-ligated cDNA for the 5´-RACE and the Marathon cDNA Synthesis primer (52-mer) was used when amplifying the adaptor ligated cDNA for the 3'-RACE.

Specific PCR primers were designed (Table 6) based on the sequence of canine $GABA_B$ receptor 1a cDNA obtained in Example 2A.

TABLE 6

Primers used in PCR to amplify the 5'- and 3'-ends of canine $GABA_B$ receptor 1a cDNA

| Nr. | Species | Sequence 5´-3´ | SEQ ID NO |
|---|---|---|---|
| 936 | canine | CTACCGCGCAATGAACTCCTCGTC | 29 |
| 1076 | canine | CGAGGTGGCGTTGGGGGTCTGTGC | 30 |
| AP1 | Marathon kit | CCATCCTAATACGACTCACTATAGGGC | 31 |
| AP2 | Marathon kit | ACTCACTATAGGGCTCGAGCGGC | 32 |

cDNA fragments encoding parts of the canine $GABA_B$ receptor were amplified by PCR from the adaptor-ligated cDNA using the designed primers. A number of different PCR programs were tested to find conditions where PCR products corresponding to $GABA_B$ receptor DNA were obtained. The 5´- PCR experiments were carried out using the Expand Long Template PCR System (Boehringer Mannheim GmbH, Germany) with the following PCR program: +94°C for 30 seconds, +72°C for 3 minutes, repeated 4 times; +94°C for 30 seconds, +70°C for 3 minutes, repeated 4 times; and +94°C for 30 seconds, +68°C for 3minutes, repeated 24 times. The primer combination AP2 and 1076 gave a PCR product which corresponded to the 5´-end of the $GABA_B$ receptor cDNA, including 114 base pairs upstream the initiation codon.

The 3´- PCR experiments were carried out using the Expand Long Template PCR System (Boehringer Mannheim GmbH, Germany) with the following PCR program: +94°C for 1 minute; and +94°C for 30 seconds, +60°C for 30 seconds,+68 °C for 4 minutes, repeated 29 times. The primer combination AP1 and 936 gave a PCR fragment which corresponded to the 3´-end of the $GABA_B$ receptor cDNA, including the poly(A) tail.

The PCR products were subcloned into the pGEM-T vector from Promega (Madison, USA). The inserts were subjected to nucleotide sequence analysis, the complete nucleotide sequence for all subclones were determined using a Thermo Sequenase dye terminator cycle sequencing pre-mix kit(Amersham Pharmacia Biotech, Uppsala, Sweden). As primers for sequencing reactions specific oligonucleotides complementary to the vector pGEM-T or primers complementary to $GABA_B$ receptor DNA were used.

Complete cDNA sequence encoding the canine $GABA_B$ receptor 1a (SEQ ID NO: 52) was obtained by aligning the sequences of the different fragments obtained in Example 2A and Example 2B.

EXAMPLE 3A: Cloning of cDNA encoding human $GABA_B$ receptor 1c and 1d from Jurkat cells

A guanidinisothiocyanate/CsCl method was used to isolate total RNA from Jurkat cells. The first-strand cDNA synthesis was performed using the First-strand cDNA Synthesis kit from Amersham Pharmacia Biotech (Uppsala, Sweden). The $pd(N)_6$ primer was used to prime the first strand synthesis. The generated cDNA molecules were used as templates in the PCR reaction described below.

Specific PCR primers (Table 7) were designed based on the sequence of human $GABA_B$ receptor 1a and 1b cDNA (Example 1), rat $GABA_B$ receptor (Kaupmann et al. 1997) and the EST sequence EMBL accession number Y11044.

TABLE 7

Primers used in RT-PCR on mRNA from Jurkat cells

| Nr | Species | Sequence 5´-3´ | SEQ ID NO |
|---|---|---|---|
| 938 | human | GACGCTTATCGAGCAGCTTC | 33 |
| 972 | human | AGCCCAGAACTCACAGGGGGACAT | 34 |
| 973 | human | GCTTCAAGCCAGGTACGAACTAA | 35 |
| 893 | rat | GGAGCACCCCCAAGCCCCACTG | 36 |
| 937 | human | CTGGTTCCTCCCAATGTG | 37 |
| 1005 | rat | CCTCTCACTCCCCTCATCTC | 38 |
| 1030 | human | AAGCCAACCTTCCCTGCTTCTC | 39 |

cDNA fragments encoding parts of the $GABA_B$ receptor were amplified directly by PCR using human and rat specific primers. All PCR experiments were carried out using the Perkin Elmer Taq DNA polymerase with Gene Amp (Roche Molecular Systems Inc., NJ,

USA) with the following PCR program: +95°C for 1 minute; +54°C for 1 minutes, +72°C for 3 minutes, repeated 44 times; and finally +72°C for 7 minutes.

The PCR products were subcloned into the pGEM-T vector from Promega (Madison, USA). The inserts were subjected to nucleotide sequence analysis, the complete nucleotide sequence for all subclones were determined using a Thermo Sequenase dye terminator cycle sequencing pre-mix kit (Amersham Pharmacia Biotech, Uppsala, Sweden). As primers for sequencing reactions specific oligonucleotides complementary to the vector pGEM-T or primers complementary to $GABA_B$ receptor DNA were used.

The following primer combination gave PCR products corresponding to the 3´-end of the $GABA_B$ receptor cDNA: primer pairs 938 and 972; 938 and 973. Unexpectedly both these fragments lacked 149 base pairs, resulting in a frame shift and the insertion of a new termination codon. The following primer combination gave a PCR product corresponding to the 5´-part of the $GABA_B$ receptor 1a cDNA: 893 and 937. The primer pairs 1005 and 937, 1030 and 937 gave PCR products corresponding to the 5´-part of the $GABA_B$ receptor 1b cDNA. Also these PCR fragments lacked the same 149 base pairs which resulted in a frame shift and the insertion of a new termination codon.

These results show that Jurkat cells contain mRNA encoding two new forms of the human $GABA_B$ receptor. These are designated $GABA_B$ receptor 1c (SEQ ID NO: 54 and 55) (comprising the same 5'-part as the $GABA_B$ receptor 1a) and $GABA_B$ receptor 1d (SEQ ID NO: 56 and 57) (comprising the same 5'-part as the $GABA_B$ receptor 1b). These two forms of the $GABA_B$ receptor do not contain any of the transmembrane regions and are therefore expected to be soluble forms of the receptor.

EXAMPLE 3B: Analysis of cDNA encoding human $GABA_B$ receptors from hippocampus

Messenger RNA from human hippocampus was obtained from Clontech (Palo Alto, USA) (catalogue #6578-1). First strand cDNA synthesis reactions were performed using the First-strand cDNA Synthesis kit from Amersham Pharmacia Biotech (Uppsala, Sweden). The

pd(N)$_6$ primer was used to prime the first-strand synthesis. The generated cDNA molecules were used as template in the PCR reactions described below.

Specific PCR primer was designed (Table 8) based on the sequences of the cDNAs encoding human GABA$_B$ receptors 1a and 1b.

TABLE 8

Primers used for RT-PCR on mRNA from human hippocampus

| Nr | Species | Sequence 5´ - 3´ | SEQ ID NO |
|---|---|---|---|
| 937 | Human | CTGGTTCCTCCCAATGTG | 40 |
| 938 | Human | GACGCTTATCGAGCAGCTTC | 41 |

cDNA fragments encoding parts of the human GABA$_B$ receptors were amplified directly by PCR using the designed primers with the generated cDNA molecules as templates. All PCR experiments were carried out using the Perkin Elmer Taq DNA polymerase with Gene Amp (Roche Molecular Systems Inc., NJ, USA) with the following PCR program: +94°C for 1 minute, +50°C for 30 seconds,+72°C for 3 minutes; +94°C for 1 minute, +54°C for 30 seconds, +72°C for 3 minutes repeated 44 times; and finally +72°C for 7 minutes. The primer combination 938 and 937 gave a PCR product which corresponded to the expected size of the GABA$_B$ receptor 1a and 1b cDNA but also a fragment a smaller size.

The PCR products were subcloned into the pGEM-T vector from Promega (Madison, USA). The inserts were subjected to nucleotide sequence analysis, the complete nucleotide sequence for all subclones were determined using a Thermo Sequenase dye terminator cycle sequencing pre-mix kit (Amersham Pharmacia Biotech, Uppsala, Sweden). As primers for sequencing reactions specific oligonucleotides complementary to the vector pGEM-T or primers complementary to the cDNA encoding the GABA$_B$ receptor were used.

The larger PCR fragment was found to correspond to the 3'-part of the GABA$_B$ receptor 1a and 1b cDNA and the smaller fragment which lacked 149 base pairs was found to correspond to the 3'-part of the GABA$_B$ receptors 1c and 1d cDNA identified in Example 3A.

EXAMPLE 4: Cloning and sequencing of cDNA encoding canine $GABA_B$ receptor 1b

cDNA encoding the canine $GABA_B$ receptor 1b is isolated in a similar manner as described in Example 2 for receptor 1a. PCR primers specifically designed to be complementary to the 5'-end of the cDNA encoding the rat and human $GABA_B$ receptor 1b, together with PCR primers complementary to the 3'-end of the cDNA encoding the canine $GABA_B$ receptor 1a, and mRNA prepared from a suitable canine tissue, are used.

EXAMPLE 5: Cloning of cDNA encoding canine $GABA_B$ receptor 1c

Total RNA from canine liver was prepared using the RNeasy Total RNA Purification Protocols (Quiagen GmbH, Germany). The first-strand cDNA synthesis was performed using the First-strand cDNA Synthesis kit from (Amersham Pharmacia Biotech, Uppsala, Sweden). The $pd(N)_6$ primer was used to prime the first -strand synthesis. The generated cDNA molecules were used as templates in the PCR reaction described below.

Specific PCR primers (Table 9) were designed based on the sequence of canine $GABA_B$ receptor 1a cDNA.

TABLE 9

Primers used in RT-PCR

| Nr | Species | Sequence 5´-3´ | SEQ ID NO |
|---|---|---|---|
| 936 | canine | CTACCGCGCAATGAACTCCTCGTC | 42 |
| 954 | canine | CCTTCTTCTCCTCCTTCTTAGTGA | 43 |

cDNA fragments encoding parts of the canine $GABA_B$ receptor were amplified directly by PCR using canine specific primers. All PCR experiments were carried out using the Perkin Elmer Taq DNA polymerase with Gene Amp (Roche Molecular Systems Inc., NJ, USA) with the following PCR program: +95°C for 1 minute, +54°C for 30 seconds, +72°C for 3

minutes, repeated 44 times and finally +72°C for 7 minutes. The primer combination gave a PCR product with a size corresponding to the $GABA_B$ receptor 1a but also a fragment of smaller size indicating the presence of a $GABA_B$ receptor 1c.

The PCR products were subcloned into the pGEM-T vector from Promega (Madison, USA). The inserts were subjected to nucleotide sequence analysis, the complete nucleotide sequence for all subclones were determined using a Thermo Sequenase dye terminator cycle sequencing pre-mix kit (Amersham Pharmacia Biotech, Uppsala, Sweden). As primers for sequencing reactions specific oligonucleotides complementary to the vector pGEM-T were used.

The smaller fragment was shown to lack 149 base pairs. This deletion caused a frame shift and an insertion of a new termination codon, verifying the existence of a canine $GABA_B$ receptor 1c.

Complete cDNA sequence encoding the canine $GABA_B$ receptor 1c (SEQ ID NO: 58) was obtained by aligning the sequences of the fragments obtained in Example 2A, Example 2B and Example 5.

EXAMPLE 6. Cloning, sequencing and organization of human $GABA_B$ receptor genomic fragments

To determine the structural organization and sequence of the human $GABA_B$ receptor gene, human genomic DNA libraries and human genomic DNA were screened and analyzed. Human genomic libraries were obtained from Clontech (Palo Alto, CA, USA). The libraries were constructed from female leukocyte DNA (catalog # HL1111J), cloned into λEMBL-3 vector. The average size of inserts are 16 kb and the number of independent clones are $1.7x10^6$. Human genomic DNA was obtained from Clontech (catalog # 6550-1). In order to isolate recombinant phages containing exon and intron sequences of the human $GABA_B$ receptor gene, 48 individual bacterial plates with a diameter of 150 mm and approximately $4 \times 10^4$ individual plaques per plate, were screened. The methods and

solutions used were as described in the Library Protocol Handbook: General Procedures for the Hybridization of Lambda Phage Libraries w/DNA Probes (Clontech) with some modifications as will be apparent from the following.

The experiment was carried out essentially as follows. The numbers are given per plate basis. A sample of the phage library diluted in 0.1 ml sterile lambda diluent was prepared in order to obtain an estimated titer of 40,000 pfu (plaque forming units). A 0.6 ml LB-medium culture of the *E. coli* host strain K802 (obtained from Clontech) was infected with 40000 pfu recombinant phages for 15 minutes at +37°C. The culture was then mixed with 7 ml top agarose (6.5 g of agarose added per liter LB) and poured onto LB plates. The plates were incubated at +37°C for approximately 7 hours. The plates were then chilled at +4°C.

Plaque hybridization experiments were as follows. Membrane filters, Colony/Plaque Screen (DuPont, Wilmington, DE, USA), were placed onto the top of the plates for 3 minutes. For denaturation of DNA the filters were removed and floated in 0.5 M NaOH on a plastic wrap for 2 minutes, with the plaque side up. This step was repeated once to ensure efficient denaturation. Following neutralization the membrane filters were placed in 1M Tris-HCl pH 7.5, two times 2 minutes and allowed to dry.

To obtain probes for DNA hybridization screening of the membrane filters, a $GABA_B$ receptor cDNA clone was digested with SacII and a 479 bp fragment (base pairs 573-1051 in the cDNA encoding human $GABA_B$ receptor 1a, SEQ ID NO: 48) was separated by agarose electrophoresis, excised and transferred to a polypropylene microcentrifuge tube. Additional probes were obtained by PCR amplification of various regions of the $GABA_B$ receptor cDNA (base pairs 68-486 and 2368-2863 in the cDNA encoding human $GABA_B$ receptor 1a, SEQ ID NO: 48). The isolated cDNA fragment was $^{32}P$-labeled using Megaprime DNA labeling system (Amersham Pharmacia Biotech, Uppsala, Sweden) by the following procedure. Water was added at a ratio of 3 ml per gram of gel, and placed in a boiling water bath for 7 minutes to melt the gel and denature the DNA. A volume of DNA/agarose solution containing 25 ng of DNA was added to the labeling reaction, according to the supplier's instructions. Labeled nucleotides were removed from DNA

labeling reactions using MicroSpin™ G-50 Columns (Amersham Pharmacia Biotech, Uppsala. Sweden ).

The DNA hybridization reaction was performed under stringent conditions according to the method described below. The filter membranes were prehybridized at +65°C for at least 1 hour in a solution composed of 1% SDS, 1M NaCl, and 10% dextran sulfate using a hybridization oven (Hybaid Ltd, Ashford, UK). Following prehybridization a solution containing denatured herring sperm DNA of a final concentration of 100 µg/ml and the $^{32}$P-labeled DNA probe at a concentration <10 ng/ml (for optimal signal to background ratio) was added to the prehybridization solution and the membrane filters were incubated at +65°C for 10-20 hours. Following the removal of the hybridization solution the membrane filters were first washed in a 2xSSC (0.3M NaCl, 0.03M Na-citrate), 1% SDS solution two times for 5 minutes at room temperature. In the next step, the membrane filters were incubated +60°C two times for 30 minutes each in the same solution. In a third step, the filters were washed two times at room temperature in 0.1xSSC. Finally, the membrane filters were placed on a sheet of filter paper with the DNA face up, and allowed to dry. The dried membrane filters were then exposed to X-ray films and autoradiographed.

Of the approximately $2x10^6$ individual plaques analyzed, four hybridizing plaques were detected and isolated. These three isolates were designated #GR1, #GR12, #GR13 and #GR41, respectively. After several rescreening experiments, the recombinant phage DNA was purified using Qiagen Lambda Midi Kit (Qiagen GmbH, Germany). The purified DNA was digested with SalI and the fragments representing the inserts were isolated by agarose electrophoresis.

The sizes of the inserts were approximately isolate #GR1, 12 kb, isolate #GR12, 12 kb; isolate #GR13, 16kb; and isolate #GR41, 19 kb. These fragments were cloned into SalI digested linearized pUC19, resulting in the plasmids pAM362 (isolate #GR1), pAM363 (isolate #GR12), pAM364 (isolate #GR13)and pAM365 (isolate #GR41). The inserts from the four plaques which hybridizes to the $GABA_B$ receptor cDNA probes were analyzed by PCR, restriction mapping and hybridization to $^{32}$P-labeled DNA fragments representing various regions of the $GABA_B$ receptor gene.

The cloned fragments in the plasmids pAM362, pAM363, pAM364 and pAM365 were characterized by restriction enzyme mapping, using EcoRI, HindIII, PstI, and BamHI. The approximate positions of the exons and the approximate size of the introns were analyzed and determined by PCR-based exon-exon linking and agarose gel electrophoresis.

In order to facilitate nucleotide sequence analysis, 7 restriction sub-fragments derived from pAM364, 2 restriction fragments derived from pAM362, and 1 restriction sub-fragment derived from pAM365 were isolated and cloned into pUC19, resulting in the plasmids pAM366-pAM375. The following strategy was employed; by combining PCR primers located within the pUC19 sequence either upstream or downstream of the cloning site, with a PCR primer with defined orientation and specific for the $GABA_B$ receptor derived subcloned fragment allowed the determination.

The inserts in the 10 plasmids pAM366-pAM375 were subjected to nucleotide sequence analysis. The nucleotide sequences for all subclones were determined using a Thermo Sequenase dye terminator cycle sequencing pre-mix kit (Amersham Pharmacia Biotech, Uppsala, Sweden). As primers for sequencing reactions specific oligonucleotides complementary to pUC19 or primers complementary to the $GABA_B$ receptor cDNA were used.

The genomic fragments cloned in the plasmids pAM362-pAM365 where shown to contain the complete transcribed part of the human $GABA_B$ receptor gene and extends > 3kb upstream of the first exon and > 2kb downstream of the last exon. The fragments cloned in the plasmid pAM362 were shown to contain exons 7-11, pAM363 exons 12-22, pAM364 exons 1-11, and pAM365 exons 12-23 of the $GABA_B$ receptor gene (Figure 1). The sequence of exons 1-11 and introns 1-10 is listed in SEQ ID NO: 60 and the sequence of exons 12-23 and introns 12-22 is listed in SEQ ID NO: 61.

The human $GABA_B$ receptor gene consists of 23 exons and 22 introns (Figure 1). The size of the exons range between 21 bp and 1486 bp. As can be noted from Table 10 exon/intron boundaries are in accordance with the AG/GT rule and conform well to the consensus sequence suggested by Mount et al. 1982.

TABLE 10

Exon-Intron boundaries of the $GABA_B$ receptor gene, sequences at exon-intron junctions.

| | | 5´splice donor | 3´splice acceptor | |
|---|---|---|---|---|
| Exon 1-Exon 2 | CGAG | GTAAGAG | CCGCCTCTCACTTAG | ATGT |
| Exon 2-Exon 3 | GAAG | GTGCATC | CGACTCACCCCTTAG | GTTG |
| Exon 3-Exon 4 | TGTG | GTGAGTA | CCWATCTCTCCACAG | TCCG |
| Exon 4-Exon 5 | CAGG | GTGAGGG | CTTTCCTGCTGCCAG | TGAA |
| Exon 5-Exon 6 | TCAG | GTGAGAT | CGCACCCCTCCTCAG | AACG |
| Exon 6-Exon 7 | CAAG | GTAGCCC | CCTCTTGTCTTTCAG | TGTG |
| Exon 7-Exon 8 | TGTG | GTAAGCA | CTCCCTGCCCCACAG | CTTT |
| Exon 8-Exon 9 | TTCG | GTGAGGA | TTATTCCCACCCAAG | ACTC |
| Exon 9-Exon10 | GAAG | GTCAGAT | CTTTCTCTGTKGTAG | CGCC |
| Exon10-Exon11 | TGAG | GTGGART | CTCCTCTGTATTCAG | GTGT |
| Exon11-Exon12 | CATG | GTGAGAG | TTTTTTCCTCCCAAG | ACAT |
| Exon12-Exon13 | CTCT | GTGAGTT | TGTTCCTTCCCTCAG | GGCC |
| Exon13-Exon14 | CAGG | GTTAGTA | TTGTCGTCTGCCCAG | GTGG |
| Exon14-Exon15 | ATTG | GTGAGTG | CCCTGTGCCATGCAG | GAGG |
| Exon15-Exon16 | TCCG | GTXAGTT | CCACCTCTGCCCTAG | TTAT |
| Exon16-Exon17 | CCAG | GTGAGGA | TCTCTTCCTTTCTAG | GCCC |
| Exon17-Exon18 | GAAG | GTGAGCT | CACATATTTATCCAG | ACTC |
| Exon18-Exon19 | TGAG | GTACCAC | TYGTTTCTGCCCTAG | ACAT |
| Exon19-Exon20 | CTTG | GTGTGTG | CTCCTGCCATCCTAG | GCAT |
| Exon20-Exon21 | GGCA | GTGAGCA | TGTCTTTCCCTCTAG | GTCC |
| Exon21-Exon22 | CAAG | GTAAGGA | AACATTTGCCCCCAG | ATGC |
| Exon22-Exon23 | TGAG | GTGCGGG | TGCTTCTTCCTCCAG | AAAG |

A comparison of the sequences of the different forms of the human $GABA_B$ receptor cDNA (SEQ ID NO: 48, 50, 54, and 56) and the sequence of the human $GABA_B$ gene (SEQ ID NO: 60 and 61) reveals that different mRNA encoding human $GABA_B$ receptors are formed by alternative splicing. The translational start of the $GABA_B$ receptor 1a is localized in exon 2 and the translational stop is localized in exon 23. The mRNA encoding $GABA_B$ receptor 1b is formed by alternative splicing where part of intron 5 is retained together with exon 6 where the translational start of the $GABA_B$ receptor 1b is derived from the intron sequence. The mRNA encoding $GABA_B$ receptor 1c is formed by alternative splicing where exon 15 are removed together with introns 14 and 15 and a frame shift and a translational stop are generated in the sequence corresponding to exon 16.

The mRNA encoding $GABA_B$ receptor 1d is formed when the translational start of the $GABA_B$ receptor 1b is generated together with the translational stop of the $GABA_B$ receptor 1c.

The 1a isoform is formed by splicing exon 5 to a cryptic splice site in the middle of exon 6. Transcription of 1b isoform mRNA is most likely initiated from regulatory elements in intron 5. The ATG initiating translation of 1b mRNA is located in the 5'-end of exon 6.

Additional variants of the mRNA encoding variants of the human $GABA_B$ receptor can be derived by alternative splicing where one ore more of the exons or part of exons are excised in the processing of the pre-mRNA. The subsequent transcription of these mRNA will give rise to variants of the human $GABA_B$ receptor with potentially different biological and/or pharmacological activities.

EXAMPLE 7. Analysis of cDNA encoding human $GABA_B$ receptors from human brain

Messenger RNA from human fetal brain (catalog #6525-1) and adult human brain (catalog #6516-1) were obtained from Clontech (Palo Alto, CA, USA). First strand cDNA synthesis reactions were performed using the first strand cDNA Synthesis kit from Amersham Pharmacia Biotech (Uppsala, Sweden).The $pd(N)_6$ primer was used to prime the first-strand synthesis. The generated cDNA molecules were used as templates in the PCR reactions described below.

Specific PCR primers was designed (Table 11 ) based on the sequences of the rat $GABA_B$ receptor 1a and 1b cDNA and human $GABA_B$ receptor 1a and 1b cDNA

TABLE 11

Primers used for RT-PCR on mRNA from human fetal brain

| Nr | Species | Sequence 5´-3´ | SEQ ID NO |
|---|---|---|---|
| 838 | Rat | ATGCGCGCCGGCAGCCAACATGCTGCTGCTGCTGCTGGTGCCTCT CTTCC | 62 |
| 863 | Rat | GGTCATCCAGCGTTGAGGTGAAGAC | 63 |
| 864 | Rat | GAAGGTTGCCAGATTATACATCCGC | 64 |
| 865 | Rat | CCACGATGATTCGAGCATCTTGACG | 65 |
| 937 | Human | CTGGTTCCTCCCAATGTG | 66 |
| 1015 | Human | CCAGTGGACTATGAGATTGAG | 67 |

cDNA fragments encoding parts of the human $GABA_B$ receptors were amplified directly by PCR using the designed primers with the generated cDNA molecules as templates. PCR experiments with primers 838, 863 ,864 and 865 were carried out using the Perkin Elmer Taq DNA polymerase with Gene Amp (Roche Molecular System Inc., NJ, USA) with the following PCR program: +95°C for 1 minute, +50°C for 30 seconds,+72°C for 3 minutes, repeated 44 times and finally +72°C for 7 minutes. PCR experiments with primers 937 and 1015 were carried out using the Expand Long Template PCR System (Boehringer Mannheim GmbH, Germany) with the following PCR program: +94°C for 2 minutes; +94°C for 10 seconds, +55°C for 30 seconds, +68°C for 3 minutes repeated 35 times; and finally +68°C for 7 minutes. The primer combinations 838 and 863, 864 and 863 , 864 and 865, 937 and 1015 gave PCR products.

The PCR products were subcloned into the pGEM-T vector from Promega (Madison, WI, USA). The inserts were subjected to nucleotide sequence analysis. The complete nucleotide sequence for all subclones were determined using a ThermoSequenase dye terminator cycle sequencing pre-mix kit (Amersham Pharmacia Biotech, Uppsala, Sweden). As for primers for sequencing reactions specific oligonucleotides complementary to the vector pGEM-T or primers complementary to the cDNA encoding the $GABA_B$ receptor were used.

A number of analyzed clones isolated from fetal brain were found to lack 186 base pairs, corresponding to exon 4. This alternative splicing resulted in a cDNA (SEQ ID NO: 70)

encoding a protein (SEQ ID NO: 71) comprising of 899 amino acids and designated $GABA_B$ receptor 1e.

Other clones from fetal brain lacked 368 base pairs, corresponding to exons 4, 5 and 6, resulting in a cDNA (SEQ ID NO: 72) where a frame shift and a translational stop codon are generated in the sequence corresponding to exon 7, encoding a protein (SEQ ID NO: 73) comprising only 97 amino acids designated $GABA_B$ receptor 1f.

One clone lacked 207 base pairs, corresponding to exons 4 and 5, resulting in a cDNA (SEQ ID NO: 74) encoding a protein (SEQ ID NO: 75) comprising 892 amino acids designated $GABA_B$ receptor 1g.

Another clone had two deletions, the first comprising 186 base pairs corresponding to exon 4, the second comprising 39 base pairs corresponding to part of exon 6, resulting in a cDNA (SEQ ID NO: 76) encoding a protein (SEQ ID NO: 77) comprising 886 amino acids designated $GABA_B$ receptor 1h.

Yet another clone from adult human brain was found to have a long deletion comprising 1194 base pairs corresponding to base pairs 319 - 1512 of the cDNA encoding human $GABA_B$ receptor 1a. This deletion corresponds to part of exon 4, exons 5-11, and part of exon 12. This cDNA (SEQ ID NO: 78) encodes a protein (SEQ ID NO: 79) comprising 563 amino acids designated $GABA_B$ receptor 1i.

One clone isolated from fetal brain lacked 284 base pairs corresponding to part of exon 3 and the whole exon 4, generating a frame shift and a translational stop codon in the sequence corresponding to exon 5. This cDNA (SEQ ID NO: 80) encodes a protein (SEQ ID NO: 81) comprising only 105 amino acids designated $GABA_B$ receptor 1j.

EXAMPLE 8. Analysis of cDNA encoding human $GABA_B$ receptors from Jurkat cells

A guanidinisothiocyanate/CsCl method was used to isolate total RNA from Jurkat cells. First strand cDNA synthesis reactions were performed using the first strand cDNA

Synthesis kit from Amersham Pharmacia Biotech (Uppsala, Sweden).The $pd(N)_6$ primer was used to prime the first-strand synthesis. The generated cDNA molecules were used as templates in the PCR reactions described below.

Specific PCR primers was designed (Table 12) based on the sequences of the human $GABA_B$ receptor 1a and 1b cDNA

TABLE 12

Primers used for RT-PCR on mRNA from Jurkat cells

| Nr | Species | Sequence 5´-3´ | SEQ ID NO |
|---|---|---|---|
| 937 | Human | CTGGTTCCTCCCAATGTG | 68 |
| 1015 | Human | CCAGTGGACTATGAGATTGAG | 69 |

cDNA fragments encoding parts of the human $GABA_B$ receptors were amplified directly by PCR using the designed primers with the generated cDNA molecules as templates. The PCR experiment was carried out using the Expand Long Template PCR System (Boehringer Mannheim GmbH, Germany) with the following PCR program: +94°C for 2 minutes; +94°C for 10 seconds, +55°C for 30 seconds, +68°C for 3 minutes repeated 35 times; and finally +68°C for 7 minutes. The primer combination 937 and 1015 gave a PCR product.

The PCR products were subcloned into the pGEM-T vector from Promega (Madison, WI, USA). The inserts were subjected to nucleotide sequence analysis. The complete nucleotide sequence for all subclones were determined using a ThermoSequenase dye terminator cycle sequencing pre-mix kit (Amersham Pharmacia Biotech, Uppsala, Sweden). As for primers for sequencing reactions specific oligonucleotides complementary to the vector pGEM-T or primers complementary to the cDNA encoding the $GABA_B$ receptor were used.

Two clones had two deletions, the first comprising 368 base pairs corresponding to exons 4, 5 and 6, the second comprising 151 base pairs corresponding to exon 15 where a frame shift and a translational stop codon are generated in the sequence corresponding to exon 7. This cDNA (SEQ ID NO: 82) encodes a protein (SEQ ID NO: 83) comprising only 98

amino acids designated $GABA_B$ receptor 1k which is identical to the $GABA_B$ receptor 1f described above.

Another two clones also had two deletions, the first comprising 246 base pairs corresponding to part of exon 4, exon 5 and exon 6, the second comprising 149 base pairs corresponding to exon 15, generating a frame shift and a translational stop codon in the sequence corresponding to exon 16. This cDNA (SEQ ID NO: 84) encodes a protein (SEQ ID NO: 5) comprising 496 amino acids designated $GABA_B$ receptor 1l.

Additional variants of cDNA encoding the human $GABA_B$ receptors can be identified in a similar manner using PCR primers based on the sequence of the cDNA and genomic fragments encoding the human $GABA_B$ receptors disclosed in the present application.

The biological activity of these variants of the human $GABA_B$ receptor can be evaluated by transfection of suitable host cells with expression vectors comprising the corresponding cDNA sequences and subsequent measurement of binding of labeled ligands or activation of the receptor or modulation of receptor function.

EXAMPLE 9: Generation of antibodies

Antibodies were raised in rabbits against four different BSA-conjugated 20 amino acids long synthetic peptides corresponding to selected regions of the human and canine $GABA_B$ receptor extracellular domain. Two were directed against a sequence common for $GABA_B$ receptor 1a and b (ab1 and ab2), one against a $GABA_B$ receptor 1a-specific region (a1) and one against a $GABA_B$ receptor 1b-specific sequence (b1). To allow BSA-conjugation, a cystein residue was added to the N-terminus in all peptides but a1, which contains an endogenous cystein.

Peptide a1: (SEQ ID NO: 49 amino acids 18-37)

Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile Ile His Pro Pro Trp

Peptide ab1: (SEQ ID NO: 49 amino acids 197-216 with N-terminally added Cys)

Cys - Glu Asp Val Asn Ser Arg Arg Asp Ile Leu Pro Asp Tyr Glu Leu Lys Leu Ile His His

Peptide ab2: (SEQ ID NO: 49 amino acids 271-290 with N-terminally added Cys)

Cys - Ser Pro Ala Leu Ser Asn Arg Gln Arg Phe Pro Thr Phe Phe Arg Thr His Pro Ser Ala

Peptide b1: (SEQ ID NO: 57 amino acids 30-47 with N-terminally added Cys)

Cys - Ser His Ser Pro His Leu Pro Arg Pro His Ser Arg Val Pro Pro His Pro Ser

The antibodies were purified from rabbit serum by affinity chromatography using the corresponding immobilized peptide and subsequently used to detect expression of recombinant $GABA_B$ receptor isoforms on Western blots.

EXAMPLE 10: Heterologous expression of $GABA_B$ receptor isoforms in mammalian cells

A HindIII / SalI cDNA fragment encoding the human $GABA_B$ receptor 1b isoform was cloned into a BPV (bovine papilloma virus)-based expression vector containing the mMT-1 (murine metallothioneine) promoter. Using the calcium phosphate method, murine C127 cells were co-transfected with the $GABA_B$ receptor expression construct and an expression plasmid containing a G418 resistance marker gene using the calcium phosphate method. G418 resistant clones were evaluated by Western blot analysis for expression of the approximately 100 kDa $GABA_B$ receptor 1b isoform(Figure 2). Similarly, but using Lipofectamine (Life Technologies Inc, Rockville, MD, USA) for transfection, the human $GABA_B$ receptor 1b isoform was expressed in human HEK-293 cells using the pCI-neo expression vector. The identity of the heterologously expressed receptor was verified in HEK-293 cells by Western blot analysis and radiologand binding experiments.

A cDNA fragment encoding the human $GABA_B$ receptor 1d isoform was cloned into a BPV-based expression vector containing the mMT-1 promotor. Using the calcium phosphate method, murine C127 cells were co-transfected with the $GABA_B$ receptor

expression construct and an expression plasmid containing a G418 resistance marker gene. G418 resistant clones, and concentrated medium from the same clones, were evaluated for $GABA_B$ receptor 1d isoform expression by Western blot analysis (Figure 3). The experiment revealed the human $GABA_B$ receptor 1d to be a secreted isoform.

EXAMPLE 11: Heterologous expression of $GABA_B$ receptor isoforms in *E. coli*

A cDNA fragment encoding the human $GABA_B$ receptor 1d isoform was cloned into a modified pET (Pharmacia Amersham, Uppsala, Sweden) vector downstream of a STII (heat stable enterotoxin II of *E. coli*) signal peptide and followed by a thrombin clevage site and a hexahistidine tag. The expression construct was subsequently used to transform the BL21 (DE3) *E. coli* strain. Western blot analysis of IPTG-induced bacteria revealed expression of a human $GABA_B$ receptor 1d isoform of expected size(Figure 4).

In addition, the human $GABA_B$ receptor 1d isoform was successfully expressed without fusion to a bacterial signal peptide in *E. coli* strain AD494 (DE3).

EXAMPLE 12: Method for the screening of substances which are $GABA_B$ receptor antagonists or agonists

$GABA_B$ receptor expressing cells, transgenic animals or cells and tissues derived thereof, are used to screen substance libraries for antagonist or agonist activities. Screening is performed as ligand binding assays or functional assays. For screening, cells and tissues are prepared in various ways, each uniquely suited to its purpose. Ligand binding assays are performed *in vivo* or *in vitro*. Functional assays exemplified by, but not limited to, $Ca^{++}$-responses, cAMP-responses and effects on $Cl^-$ and $K^+$ channels, are performed in living cells, broken cells, isolated cell membranes, as well as in tissues and in living animals.

CLAIMS

1. A nucleic acid molecule encoding a human or canine $GABA_B$ receptor, or a functionally equivalent modified form thereof.

2. A nucleic acid molecule according to claim 1 encoding a human $GABA_B$ receptor, or a functionally equivalent modified form thereof.

3. A nucleic acid molecule according to claim 2 encoding the human $GABA_B$ receptor 1a, or a functionally equivalent modified form thereof.

4. A nucleic acid molecule according to claim 3 selected from:
(a) a nucleic acid molecule comprising a nucleotide sequence set forth as SEQ ID NO: 48;
(b) a nucleic acid molecule comprising a nucleotide sequence capable of hybridizing, under stringent hybridization conditions, to a nucleotide sequence complementary to the polypeptide coding region of a DNA molecule as defined in (a); and
(c) a nucleic acid molecule comprising a nucleotide sequence which is degenerate as a result of the genetic code to a nucleotide sequence as defined in (a) or (b).

5. A nucleic acid molecule according to claim 2 encoding the human $GABA_B$ receptor 1b, or a functionally equivalent modified form thereof..

6. A nucleic acid molecule according to claim 5 selected from:
(a) nucleic acid molecules comprising a nucleotide sequence set forth as SEQ ID NO: 50;
(b) nucleic acid molecules comprising a nucleotide sequence capable of hybridizing, under stringent hybridization conditions, to a nucleotide sequence complementary the polypeptide coding region of a DNA molecule as defined in (a); and
(c) nucleic acid molecules comprising a nucleotide sequence which is degenerate as a result of the genetic code to a nucleotide sequence as defined in (a) or (b).

7. A nucleic acid molecule according to claim 2 encoding the human $GABA_B$ receptor 1c, or a functionally equivalent modified form thereof.

8. A nucleic acid molecule according to claim 7 selected from:
(a) nucleic acid molecules comprising a nucleotide sequence set forth as SEQ ID NO: 54;
(b) nucleic acid molecules comprising a nucleotide sequence capable of hybridizing, under stringent hybridization conditions, to a nucleotide sequence complementary the polypeptide coding region of a DNA molecule as defined in (a); and
(c) nucleic acid molecules comprising a nucleotide sequence which is degenerate as a result of the genetic code to a nucleotide sequence as defined in (a) or (b).

9. A nucleic acid molecule according to claim 2 encoding the human $GABA_B$ receptor 1d, or a functionally equivalent modified form thereof..

10. A nucleic acid molecule according to claim 9 selected from:
(a) nucleic acid molecules comprising a nucleotide sequence set forth as SEQ ID NO: 56;
(b) nucleic acid molecules comprising a nucleotide sequence capable of hybridizing, under stringent hybridization conditions, to a nucleotide sequence complementary the polypeptide coding region of a DNA molecule as defined in (a); and
(c) nucleic acid molecules comprising a nucleotide sequence which is degenerate as a result of the genetic code to a nucleotide sequence as defined in (a) or (b).

11. A nucleic acid molecule according to claim 1 encoding a canine $GABA_B$ receptor, or a functionally equivalent modified form thereof.

12. A nucleic acid molecule according to claim 11 encoding the canine $GABA_B$ receptor 1a, or a functionally equivalent modified form thereof.

13. A nucleic acid molecule according to claim 12 selected from

(a) nucleic acid molecules comprising a nucleotide sequence set forth as SEQ ID NO: 52;

(b) nucleic acid molecules comprising a nucleotide sequence capable of hybridizing, under stringent hybridization conditions, to a nucleotide sequence complementary the polypeptide coding region of a DNA molecule as defined in (a); and

(c) nucleic acid molecules comprising a nucleotide sequence which is degenerate as a result of the genetic code to a nucleic acid sequence as defined in (a) or (b).

14 A nucleic acid molecule according to claim 11 encoding the canine $GABA_B$ receptor 1c, or a functionally equivalent modified form thereof.

15. A nucleic acid molecule according to claim 14 selected from

(a) nucleic acid molecules comprising a nucleotide sequence set forth as SEQ ID NO: 58;

(b) nucleic acid molecules comprising a nucleotide sequence capable of hybridizing, under stringent hybridization conditions, to a nucleotide sequence complementary the polypeptide coding region of a DNA molecule as defined in (a); and

(c) nucleic acid molecules comprising a nucleotide sequence which is degenerate as a result of the genetic code to a nucleic acid sequence as defined in (a) or (b).

16. A recombinant polypeptide encoded by a nucleotide sequence according to any one of claims 1 to 15.

17. The polypeptide according to claim 16 comprising an amino acid sequence set forth as SEQ ID NO: 49, 51, 53, 55, 57 or 59 in the Sequence Listing.

18. A polypeptide according to claim 16 which has been posttranslationally modified.

19. A vector transformed with a nucleic acid molecule according to any one of claims 1 to 15.

20. A cultured host cell harboring a vector according to claim 19.

21. A process for the production of a $GABA_B$ receptor, said process comprising culturing a host cell according to claim 20 under conditions whereby the said polypeptide is produced, and recovering the said polypeptide.

22. A method for the screening of compounds which are inhibitors of transient lower esophageal sphincter relaxations, said method comprising the use of a nucleic acid molecule encoding a $GABA_B$ receptor.

23. A method according to claim 22 wherein the said nucleic acid molecule encoding a $GABA_B$ receptor is the nucleic acid molecule according to any one of claims 1 to 15.

24. A method for the screening of compounds which are agonists or antagonists to a $GABA_B$ receptor, said method comprising the use of a nucleic acid molecule according to any one of claims 1 to 15.

25. A method according to any one of claims 22 to 24 comprising the steps (a) transforming a cultured cell with a nucleic acid molecule encoding a $GABA_B$ receptor, so that a $GABA_B$ receptor is expressed on the surface of the cell; (b) contacting a test compound with the said cell; and (c) determining whether the test compound binds to, and/or activate, the $GABA_B$ receptor.

26. A pharmaceutical compositions comprising a soluble $GABA_B$ receptor further comprising one or more pharamceutical acceptable carriers and/or diluents.

1/2

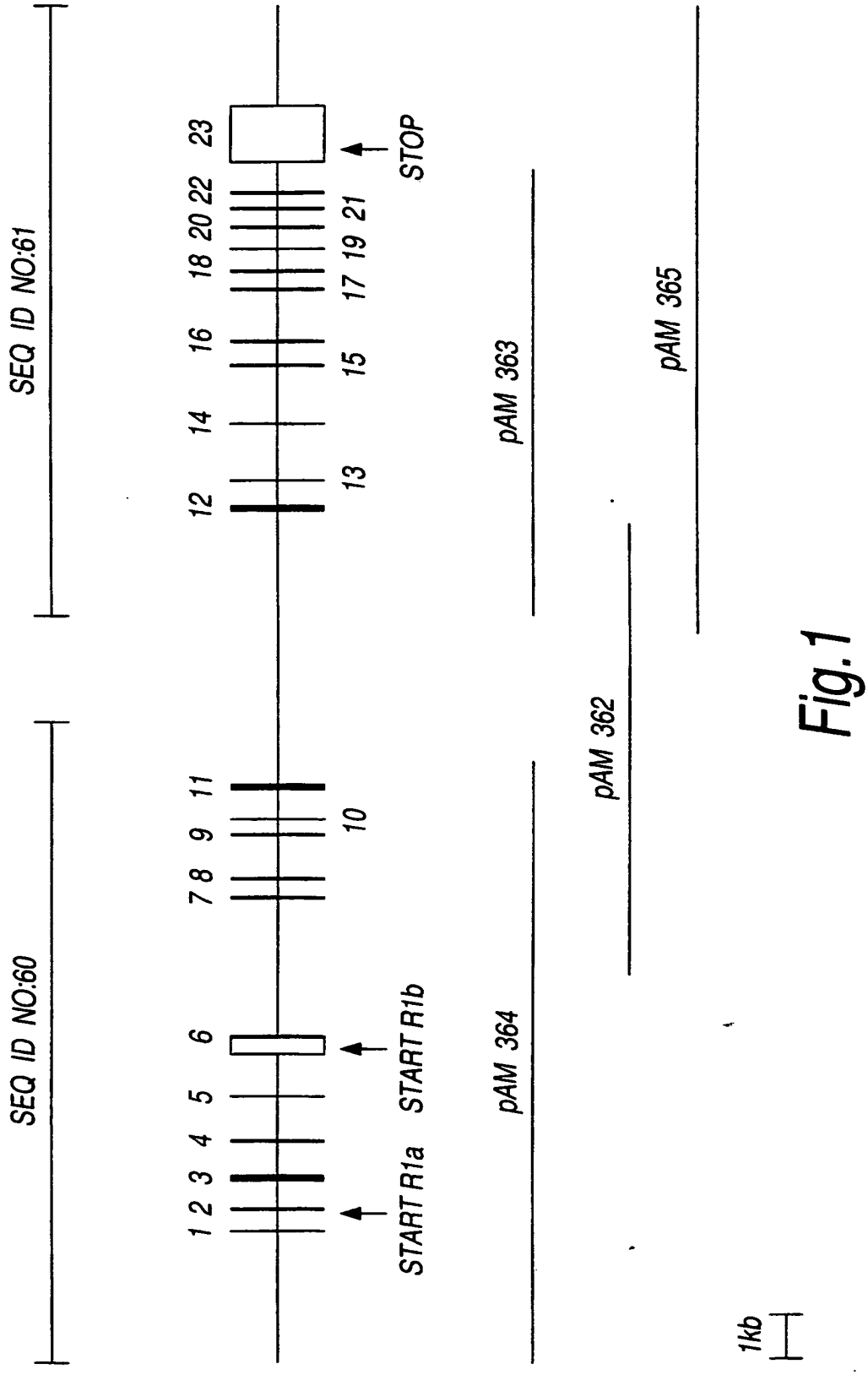
SEQ ID NO:61
SEQ ID NO:60
1 2 3 4 5 6 7 8 9 11 10 12 13 14 16 15 17 19 18 20 22 21 23
START R1a
START R1b
STOP
pAM 364
pAM 362
pAM 363
pAM 365
1kb

Fig.1

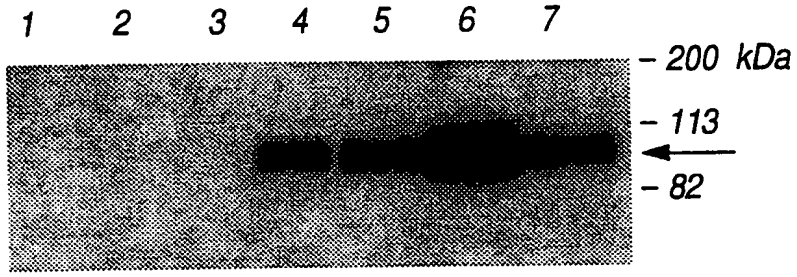


Fig.2

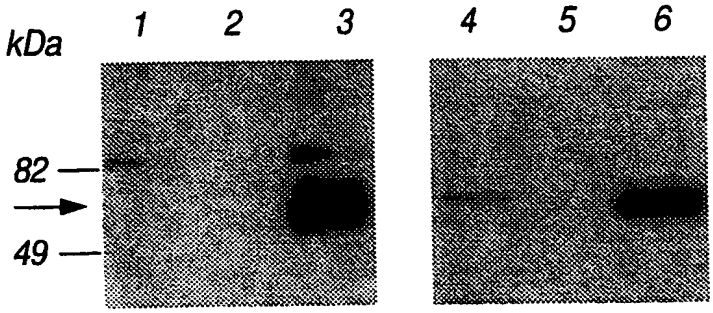


Fig.3

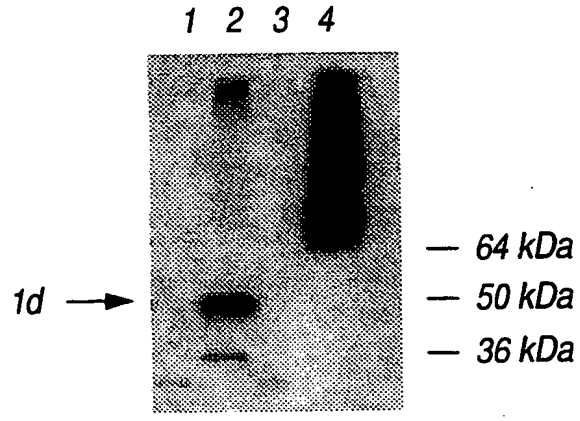


Fig.4

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: ASTRA AB
(B) STREET: Västra Mälarehamnen 9
(C) CITY: Södertälje
(E) COUNTRY: Sweden
(F) POSTAL CODE (ZIP): S-151 85
(G) TELEPHONE: +46-8-553 260 00
(H) TELEFAX: +46-8-553 288 20
(I) TELEX: 19237 astra s

(ii) TITLE OF INVENTION: New nucleotide sequences

(iii) NUMBER OF SEQUENCES: 85

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GTTTCTTCTC GGATCCAGCT GTGCCTG 27

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

CAGGCACAGC TGGATCCGAG AAGAAACT 28

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

CGGTCGACTC ACTTGTAAAG CAAATGTACT CGACTCCC 38

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

ATGCGCGCCG GCAGCCAACA TGCTGCTGCT GCTGCTGGTG C 41

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CGGTCGACTC ACTTGTAAAG CAAATGTACT CGACTCCCAT CACAGC 46

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

ATGCGCGCCG GCAGCCAACA TGCTGCTGCT GCTGCTGGTG CCTCTCTTCC 50

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

CAGGCACAGC TGGATCCGAG AAGAAACTCT GTCGGAAAGT 40

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GGTCATCCAG CGTTGAGGTG AAGAC 25

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GAAGGTTGCC AGATTATACA TCCGC 25

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CCACGATGAT TCGAGCATCT TGACG 25

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GCCTCTCACT CCCCTCATCT CC 22

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GAGTGAAGGA GGCTGGAATT G 21

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GACGCTTATC GAGCAGCTTC 20

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

AGCCCAGAAC TCACAGGGGG ACAT 24

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GCTTCAAGCC AGGTACGAAC TAA 23

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

TGGCCCTCCA CCGCCTCAGT CATCTCA 27

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

CCATCCTAAT ACGACTCACT ATAGGGC 27

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

CTCAATCTCA TAGTCCACTG G 21

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

CCTTGAGGCC CGGGGAGAG 19

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

CAGGCACAGC TGGATCCGAG AAGAAACT 28

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

CGGTCGACTC ACTTGTAAAG CAAATGTACT CGACTCCCAT CACAGC 46

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

ATGCGCGCCG GCAGCCAACA TGCTGCTGCT GCTGCTGGTG CCTCTCTTCC 50

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

CGTCAAGATG CTCGAATCAT CG 22

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

CAGGGGGCTC AGAGGGTCCC 20

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

CAGGCACAGC TGGATCCGAG AAGAAACTCT GTCGGAAAGT 40

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

CGGTCGACTC ACTTGTAAAG CAAATGTACT CGACTCCCAT CACAGCTAAG 50

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

ACTTTCCGAC AGAGTTTCTT CTCGGATCCA GCTGTGCCTG 40

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

CCACGATGAT TCGAGCATCT TGACG 25

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

CTACCGCGCA ATGAACTCCT CGTC 24

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

CGAGGTGGCG TTGGGGGTCT GTGC 24

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

CCATCCTAAT ACGACTCACT ATAGGGC 27

(2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

ACTCACTATA GGGCTCGAGC GGC 23

(2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

GACGCTTATC GAGCAGCTTC 20

(2) INFORMATION FOR SEQ ID NO: 34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

AGCCCAGAAC TCACAGGGGG ACAT 24

(2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

GCTTCAAGCC AGGTACGAAC TAA 23

(2) INFORMATION FOR SEQ ID NO: 36:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

GGAGCACCCC CAAGCCCCAC TG 22

(2) INFORMATION FOR SEQ ID NO: 37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

CTGGTTCCTC CCAATGTG 18

(2) INFORMATION FOR SEQ ID NO: 38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

CCTCTCACTC CCCTCATCTC 20

(2) INFORMATION FOR SEQ ID NO: 39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

AAGCCAACCT TCCCTGCTTC TC 22

(2) INFORMATION FOR SEQ ID NO: 40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

CTGGTTCCTC CCAATGTG 18

(2) INFORMATION FOR SEQ ID NO: 41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

GACGCTTATC GAGCAGCTTC 20

(2) INFORMATION FOR SEQ ID NO: 42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

CTACCGCGCA ATGAACTCCT CGTC 24

(2) INFORMATION FOR SEQ ID NO: 43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

CCTTCTTCTC CTCCTTCTTA GTGA 24

(2) INFORMATION FOR SEQ ID NO: 44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2883 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Rattus norwegicus

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..2883

(x) PUBLICATION INFORMATION:
(A) AUTHORS: Kaupmann, K
Huggel, K
Heid, J
Flor, P
Bischoff, S
Mickel, S
McMaster, G
Angst, C
Bittiger, H
Froestl, W
(B) TITLE: Expression cloning of GABA-B receptors uncovers similarity to metabotropic glutamate receptors
(C) JOURNAL: Nature
(D) VOLUME: 386
(F) PAGES: 239-246
(G) DATE: 20 march-1997

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

```
ATG CTG CTG CTG CTG CTG GTG CCT CTC TTC CTC CGC CCC CTG GGC GCT      48
Met Leu Leu Leu Leu Leu Val Pro Leu Phe Leu Arg Pro Leu Gly Ala
 1               5                  10                  15

GGC GGG GCG CAG ACC CCC AAC GCC ACC TCG GAA GGT TGC CAG ATT ATA      96
Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile Ile
            20                  25                  30

CAT CCG CCC TGG GAA GGT GGC ATC AGG TAC CGT GGC TTG ACT CGC GAC     144
His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg Asp
        35                  40                  45

CAG GTG AAG GCC ATC AAC TTC CTG CCT GTG GAC TAT GAG ATC GAA TAT     192
```

```
Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu Tyr
     50                  55                  60

GTG TGC CGA GGG GAG CGC GAG GTG GTG GGG CCC AAG GTG CGC AAA TGC      240
Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys Cys
 65                  70                  75                  80

CTG GCC AAC GGC TCC TGG ACG GAT ATG GAC ACA CCC AGC CGC TGT GTC      288
Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys Val
                 85                  90                  95

CGA ATC TGC TCC AAG TCT TAT TTG ACC CTG GAA AAT GGG AAG GTT TTC      336
Arg Ile Cys Ser Lys Ser Tyr Leu Thr Leu Glu Asn Gly Lys Val Phe
            100                 105                 110

CTG ACG GGT GGG GAC CTC CCA GCT CTG GAT GGA GCC CGG GTG GAG TTC      384
Leu Thr Gly Gly Asp Leu Pro Ala Leu Asp Gly Ala Arg Val Glu Phe
        115                 120                 125

CGA TGT GAC CCC GAC TTC CAT CTG GTG GGC AGC TCC CGG AGC GTC TGT      432
Arg Cys Asp Pro Asp Phe His Leu Val Gly Ser Ser Arg Ser Val Cys
    130                 135                 140

AGT CAG GGC CAG TGG AGC ACC CCC AAG CCC CAC TGC CAG GTG AAT CGA      480
Ser Gln Gly Gln Trp Ser Thr Pro Lys Pro His Cys Gln Val Asn Arg
145                 150                 155                 160

ACG CCA CAC TCA GAA CGG CGT GCA GTA TAC ATC GGG GCG CTG TTT CCC      528
Thr Pro His Ser Glu Arg Arg Ala Val Tyr Ile Gly Ala Leu Phe Pro
                165                 170                 175

ATG AGC GGG GGC TGG CCG GGG GGC CAG GCC TGC CAG CCC GCG GTG GAG      576
Met Ser Gly Gly Trp Pro Gly Gly Gln Ala Cys Gln Pro Ala Val Glu
            180                 185                 190

ATG GCG CTG GAG GAC GTT AAC AGC CGC AGA GAC ATC CTG CCG GAC TAC      624
Met Ala Leu Glu Asp Val Asn Ser Arg Arg Asp Ile Leu Pro Asp Tyr
        195                 200                 205

GAG CTC AAG CTT ATC CAC CAC GAC AGC AAG TGT GAC CCA GGG CAA GCC      672
Glu Leu Lys Leu Ile His His Asp Ser Lys Cys Asp Pro Gly Gln Ala
    210                 215                 220
```

```
ACC AAG TAC TTG TAC GAA CTA CTC TAC AAT GAC CCC ATC AAG ATC ATT      720
Thr Lys Tyr Leu Tyr Glu Leu Leu Tyr Asn Asp Pro Ile Lys Ile Ile
225                 230                 235                 240

CTC ATG CCT GGC TGT AGT TCT GTC TCC ACA CTT GTA GCT GAG GCT GCC      768
Leu Met Pro Gly Cys Ser Ser Val Ser Thr Leu Val Ala Glu Ala Ala
                245                 250                 255

CGG ATG TGG AAC CTT ATT GTG CTC TCA TAT GGC TCC AGT TCA CCA GCC      816
Arg Met Trp Asn Leu Ile Val Leu Ser Tyr Gly Ser Ser Ser Pro Ala
            260                 265                 270

TTG TCA AAC CGA CAG CGG TTT CCC ACG TTC TTC CGG ACG CAT CCA TCC      864
Leu Ser Asn Arg Gln Arg Phe Pro Thr Phe Phe Arg Thr His Pro Ser
        275                 280                 285

GCC ACA CTC CAC AAT CCC ACC CGG GTG AAA CTC TTC GAA AAG TGG GGC      912
Ala Thr Leu His Asn Pro Thr Arg Val Lys Leu Phe Glu Lys Trp Gly
    290                 295                 300

TGG AAG AAG ATC GCT ACC ATC CAA CAG ACC ACC GAG GTC TTC ACC TCA      960
Trp Lys Lys Ile Ala Thr Ile Gln Gln Thr Thr Glu Val Phe Thr Ser
305                 310                 315                 320

ACG CTG GAT GAC CTG GAG GAG CGA GTG AAA GAG GCT GGG ATC GAG ATC     1008
Thr Leu Asp Asp Leu Glu Glu Arg Val Lys Glu Ala Gly Ile Glu Ile
                325                 330                 335

ACT TTC CGA CAG AGT TTC TTC TCG GAT CCA GCT GTG CCT GTT AAA AAC     1056
Thr Phe Arg Gln Ser Phe Phe Ser Asp Pro Ala Val Pro Val Lys Asn
            340                 345                 350

CTG AAG CGT CAA GAT GCT CGA ATC ATC GTG GGA CTT TTC TAT GAG ACG     1104
Leu Lys Arg Gln Asp Ala Arg Ile Ile Val Gly Leu Phe Tyr Glu Thr
        355                 360                 365

GAA GCC CGG AAA GTT TTT TGT GAG GTC TAT AAG GAA AGG CTC TTT GGG     1152
Glu Ala Arg Lys Val Phe Cys Glu Val Tyr Lys Glu Arg Leu Phe Gly
    370                 375                 380

AAG AAG TAC GTC TGG TTC CTC ATC GGG TGG TAT GCT GAC AAC TGG TTC     1200
Lys Lys Tyr Val Trp Phe Leu Ile Gly Trp Tyr Ala Asp Asn Trp Phe
385                 390                 395                 400
```

```
AAG ACC TAT GAC CCG TCA ATC AAT TGT ACA GTG GAA GAA ATG ACC GAG      1248
Lys Thr Tyr Asp Pro Ser Ile Asn Cys Thr Val Glu Glu Met Thr Glu
                405                 410                 415

GCG GTG GAG GGC CAC ATC ACC ACG GAG ATT GTC ATG CTG AAC CCT GCC      1296
Ala Val Glu Gly His Ile Thr Thr Glu Ile Val Met Leu Asn Pro Ala
            420                 425                 430

AAC ACC CGA AGC ATT TCC AAC ATG ACG TCA CAG GAA TTT GTG GAG AAA      1344
Asn Thr Arg Ser Ile Ser Asn Met Thr Ser Gln Glu Phe Val Glu Lys
        435                 440                 445

CTA ACC AAG CGG CTG AAA AGA CAC CCC GAG GAG ACT GGA GGC TTC CAG      1392
Leu Thr Lys Arg Leu Lys Arg His Pro Glu Glu Thr Gly Gly Phe Gln
    450                 455                 460

GAG GCA CCA CTG GCC TAT GAT GCT ATC TGG GCC TTG GCT TTG GCC TTG      1440
Glu Ala Pro Leu Ala Tyr Asp Ala Ile Trp Ala Leu Ala Leu Ala Leu
465                 470                 475                 480

AAC AAG ACG TCT GGA GGA GGT GGT CGT TCC GGC GTG CGC CTG GAG GAC      1488
Asn Lys Thr Ser Gly Gly Gly Gly Arg Ser Gly Val Arg Leu Glu Asp
                485                 490                 495

TTT AAC TAC AAC AAC CAG ACC ATT ACA GAC CAG ATC TAC CGG GCC ATG      1536
Phe Asn Tyr Asn Asn Gln Thr Ile Thr Asp Gln Ile Tyr Arg Ala Met
            500                 505                 510

AAC TCC TCC TCC TTT GAG GGC GTT TCT GGC CAT GTG GTC TTT GAT GCC      1584
Asn Ser Ser Ser Phe Glu Gly Val Ser Gly His Val Val Phe Asp Ala
        515                 520                 525

AGC GGC TCC CGG ATG GCA TGG ACA CTT ATC GAG CAG CTA CAG GGC GGC      1632
Ser Gly Ser Arg Met Ala Trp Thr Leu Ile Glu Gln Leu Gln Gly Gly
    530                 535                 540

AGC TAC AAG AAG ATC GGC TAC TAC GAC AGC ACC AAG GAT GAT CTT TCC      1680
Ser Tyr Lys Lys Ile Gly Tyr Tyr Asp Ser Thr Lys Asp Asp Leu Ser
545                 550                 555                 560

TGG TCC AAA ACG GAC AAG TGG ATT GGA GGG TCT CCC CCA GCT GAC CAG      1728
Trp Ser Lys Thr Asp Lys Trp Ile Gly Gly Ser Pro Pro Ala Asp Gln
```

```
                    565                 570                 575

ACC TTG GTC ATC AAG ACA TTC CGT TTC CTG TCT CAG AAA CTC TTT ATC      1776
Thr Leu Val Ile Lys Thr Phe Arg Phe Leu Ser Gln Lys Leu Phe Ile
                580                 585                 590

TCC GTC TCA GTT CTC TCC AGC CTG GGC ATT GTT CTT GCT GTT GTC TGT      1824
Ser Val Ser Val Leu Ser Ser Leu Gly Ile Val Leu Ala Val Val Cys
        595                 600                 605

CTG TCC TTT AAC ATC TAC AAC TCC CAC GTT CGT TAT ATC CAG AAC TCC      1872
Leu Ser Phe Asn Ile Tyr Asn Ser His Val Arg Tyr Ile Gln Asn Ser
    610                 615                 620

CAG CCC AAC CTG AAC AAT CTG ACT GCT GTG GGC TGC TCA CTG GCA CTG      1920
Gln Pro Asn Leu Asn Asn Leu Thr Ala Val Gly Cys Ser Leu Ala Leu
625                 630                 635                 640

GCT GCT GTC TTC CCT CTC GGG CTG GAT GGT TAC CAC ATA GGG AGA AGC      1968
Ala Ala Val Phe Pro Leu Gly Leu Asp Gly Tyr His Ile Gly Arg Ser
                645                 650                 655

CAG TTC CCG TTT GTC TGC CAG GCC CGC CTT TGG CTC TTG GGC TTG GGC      2016
Gln Phe Pro Phe Val Cys Gln Ala Arg Leu Trp Leu Leu Gly Leu Gly
            660                 665                 670

TTT AGT CTG GGC TAT GGC TCT ATG TTC ACC AAG ATC TGG TGG GTC CAC      2064
Phe Ser Leu Gly Tyr Gly Ser Met Phe Thr Lys Ile Trp Trp Val His
        675                 680                 685

ACA GTC TTC ACG AAG AAG GAG GAG AAG AAG GAG TGG AGG AAG ACC CTA      2112
Thr Val Phe Thr Lys Lys Glu Glu Lys Lys Glu Trp Arg Lys Thr Leu
    690                 695                 700

GAG CCC TGG AAA CTC TAT GCC ACT GTG GGC CTG CTG GTG GGC ATG GAT      2160
Glu Pro Trp Lys Leu Tyr Ala Thr Val Gly Leu Leu Val Gly Met Asp
705                 710                 715                 720

GTC CTG ACT CTT GCC ATC TGG CAG ATT GTG GAC CCC TTG CAC CGA ACC      2208
Val Leu Thr Leu Ala Ile Trp Gln Ile Val Asp Pro Leu His Arg Thr
                725                 730                 735

ATT GAG ACT TTT GCC AAG GAG GAA CCA AAG GAA GAC ATC GAT GTC TCC      2256
```

```
Ile Glu Thr Phe Ala Lys Glu Glu Pro Lys Glu Asp Ile Asp Val Ser
            740                 745                 750

ATT CTG CCC CAG TTG GAG CAC TGC AGC TCC AAG AAG ATG AAT ACG TGG      2304
Ile Leu Pro Gln Leu Glu His Cys Ser Ser Lys Lys Met Asn Thr Trp
        755                 760                 765

CTT GGC ATT TTC TAT GGT TAC AAG GGG CTG CTG CTG CTG CTG GGA ATC      2352
Leu Gly Ile Phe Tyr Gly Tyr Lys Gly Leu Leu Leu Leu Gly Ile
    770                 775                 780

TTT CTT GCT TAC GAA ACC AAG AGC GTG TCC ACT GAA AAG ATC AAT GAC      2400
Phe Leu Ala Tyr Glu Thr Lys Ser Val Ser Thr Glu Lys Ile Asn Asp
785                 790                 795                 800

CAC AGG GCC GTG GGC ATG GCT ATC TAC AAT GTC GCG GTC CTG TGT CTC      2448
His Arg Ala Val Gly Met Ala Ile Tyr Asn Val Ala Val Leu Cys Leu
                805                 810                 815

ATC ACT GCT CCT GTG ACC ATG ATC CTT TCC AGT CAG CAG GAC GCA GCC      2496
Ile Thr Ala Pro Val Thr Met Ile Leu Ser Ser Gln Gln Asp Ala Ala
            820                 825                 830

TTT GCC TTT GCC TCT CTG GCC ATC GTG TTC TCT TCC TAC ATC ACT CTG      2544
Phe Ala Phe Ala Ser Leu Ala Ile Val Phe Ser Ser Tyr Ile Thr Leu
        835                 840                 845

GTT GTG CTC TTT GTG CCC AAG ATG CGC AGG CTG ATC ACC CGA GGG GAA      2592
Val Val Leu Phe Val Pro Lys Met Arg Arg Leu Ile Thr Arg Gly Glu
    850                 855                 860

TGG CAG TCT GAA ACG CAG GAC ACC ATG AAA ACA GGA TCA TCC ACC AAC      2640
Trp Gln Ser Glu Thr Gln Asp Thr Met Lys Thr Gly Ser Ser Thr Asn
865                 870                 875                 880

AAC AAC GAG GAA GAG AAG TCC CGA CTG TTG GAG AAG GAA AAC CGA GAA      2688
Asn Asn Glu Glu Glu Lys Ser Arg Leu Leu Glu Lys Glu Asn Arg Glu
                885                 890                 895

CTG GAA AAG ATC ATC GCT GAG AAA GAG GAG CGC GTC TCT GAA CTG CGC      2736
Leu Glu Lys Ile Ile Ala Glu Lys Glu Glu Arg Val Ser Glu Leu Arg
            900                 905                 910
```

```
CAT CAG CTC CAG TCT CGG CAG CAA CTC CGC TCA CGG CGC CAC CCC CCA      2784
His Gln Leu Gln Ser Arg Gln Gln Leu Arg Ser Arg Arg His Pro Pro
        915                 920                 925

ACA CCC CCA GAT CCC TCT GGG GGC CTT CCC AGG GGA CCC TCT GAG CCC      2832
Thr Pro Pro Asp Pro Ser Gly Gly Leu Pro Arg Gly Pro Ser Glu Pro
    930                 935                 940

CCT GAC CGG CTT AGC TGT GAT GGG AGT CGA GTA CAT TTG CTT TAC AAG      2880
Pro Asp Arg Leu Ser Cys Asp Gly Ser Arg Val His Leu Leu Tyr Lys
945                 950                 955                 960

TGA                                                                 2883
 *
```

(2) INFORMATION FOR SEQ ID NO: 45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 960 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

```
Met Leu Leu Leu Leu Leu Val Pro Leu Phe Leu Arg Pro Leu Gly Ala
 1               5                  10                  15

Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile Ile
            20                  25                  30

His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg Asp
        35                  40                  45

Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu Tyr
    50                  55                  60

Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys Cys
65                  70                  75                  80

Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys Val
```

85 90 95

Arg Ile Cys Ser Lys Ser Tyr Leu Thr Leu Glu Asn Gly Lys Val Phe
100 105 110

Leu Thr Gly Gly Asp Leu Pro Ala Leu Asp Gly Ala Arg Val Glu Phe
115 120 125

Arg Cys Asp Pro Asp Phe His Leu Val Gly Ser Ser Arg Ser Val Cys
130 135 140

Ser Gln Gly Gln Trp Ser Thr Pro Lys Pro His Cys Gln Val Asn Arg
145 150 155 160

Thr Pro His Ser Glu Arg Arg Ala Val Tyr Ile Gly Ala Leu Phe Pro
165 170 175

Met Ser Gly Gly Trp Pro Gly Gly Gln Ala Cys Gln Pro Ala Val Glu
180 185 190

Met Ala Leu Glu Asp Val Asn Ser Arg Arg Asp Ile Leu Pro Asp Tyr
195 200 205

Glu Leu Lys Leu Ile His His Asp Ser Lys Cys Asp Pro Gly Gln Ala
210 215 220

Thr Lys Tyr Leu Tyr Glu Leu Leu Tyr Asn Asp Pro Ile Lys Ile Ile
225 230 235 240

Leu Met Pro Gly Cys Ser Ser Val Ser Thr Leu Val Ala Glu Ala Ala
245 250 255

Arg Met Trp Asn Leu Ile Val Leu Ser Tyr Gly Ser Ser Ser Pro Ala
260 265 270

Leu Ser Asn Arg Gln Arg Phe Pro Thr Phe Phe Arg Thr His Pro Ser
275 280 285

Ala Thr Leu His Asn Pro Thr Arg Val Lys Leu Phe Glu Lys Trp Gly
290 295 300

Trp Lys Lys Ile Ala Thr Ile Gln Gln Thr Thr Glu Val Phe Thr Ser
305 310 315 320

```
Thr Leu Asp Asp Leu Glu Glu Arg Val Lys Glu Ala Gly Ile Glu Ile
                325                 330                 335

Thr Phe Arg Gln Ser Phe Phe Ser Asp Pro Ala Val Pro Val Lys Asn
            340                 345                 350

Leu Lys Arg Gln Asp Ala Arg Ile Ile Val Gly Leu Phe Tyr Glu Thr
        355                 360                 365

Glu Ala Arg Lys Val Phe Cys Glu Val Tyr Lys Glu Arg Leu Phe Gly
    370                 375                 380

Lys Lys Tyr Val Trp Phe Leu Ile Gly Trp Tyr Ala Asp Asn Trp Phe
385                 390                 395                 400

Lys Thr Tyr Asp Pro Ser Ile Asn Cys Thr Val Glu Glu Met Thr Glu
                405                 410                 415

Ala Val Glu Gly His Ile Thr Thr Glu Ile Val Met Leu Asn Pro Ala
            420                 425                 430

Asn Thr Arg Ser Ile Ser Asn Met Thr Ser Gln Glu Phe Val Glu Lys
        435                 440                 445

Leu Thr Lys Arg Leu Lys Arg His Pro Glu Glu Thr Gly Gly Phe Gln
    450                 455                 460

Glu Ala Pro Leu Ala Tyr Asp Ala Ile Trp Ala Leu Ala Leu Ala Leu
465                 470                 475                 480

Asn Lys Thr Ser Gly Gly Gly Gly Arg Ser Gly Val Arg Leu Glu Asp
                485                 490                 495

Phe Asn Tyr Asn Asn Gln Thr Ile Thr Asp Gln Ile Tyr Arg Ala Met
            500                 505                 510

Asn Ser Ser Ser Phe Glu Gly Val Ser Gly His Val Val Phe Asp Ala
        515                 520                 525

Ser Gly Ser Arg Met Ala Trp Thr Leu Ile Glu Gln Leu Gln Gly Gly
    530                 535                 540
```

Ser Tyr Lys Lys Ile Gly Tyr Tyr Asp Ser Thr Lys Asp Asp Leu Ser
545 550 555 560

Trp Ser Lys Thr Asp Lys Trp Ile Gly Gly Ser Pro Pro Ala Asp Gln
565 570 575

Thr Leu Val Ile Lys Thr Phe Arg Phe Leu Ser Gln Lys Leu Phe Ile
580 585 590

Ser Val Ser Val Leu Ser Ser Leu Gly Ile Val Leu Ala Val Val Cys
595 600 605

Leu Ser Phe Asn Ile Tyr Asn Ser His Val Arg Tyr Ile Gln Asn Ser
610 615 620

Gln Pro Asn Leu Asn Asn Leu Thr Ala Val Gly Cys Ser Leu Ala Leu
625 630 635 640

Ala Ala Val Phe Pro Leu Gly Leu Asp Gly Tyr His Ile Gly Arg Ser
645 650 655

Gln Phe Pro Phe Val Cys Gln Ala Arg Leu Trp Leu Leu Gly Leu Gly
660 665 670

Phe Ser Leu Gly Tyr Gly Ser Met Phe Thr Lys Ile Trp Trp Val His
675 680 685

Thr Val Phe Thr Lys Lys Glu Glu Lys Lys Glu Trp Arg Lys Thr Leu
690 695 700

Glu Pro Trp Lys Leu Tyr Ala Thr Val Gly Leu Leu Val Gly Met Asp
705 710 715 720

Val Leu Thr Leu Ala Ile Trp Gln Ile Val Asp Pro Leu His Arg Thr
725 730 735

Ile Glu Thr Phe Ala Lys Glu Glu Pro Lys Glu Asp Ile Asp Val Ser
740 745 750

Ile Leu Pro Gln Leu Glu His Cys Ser Ser Lys Lys Met Asn Thr Trp
755 760 765

Leu Gly Ile Phe Tyr Gly Tyr Lys Gly Leu Leu Leu Leu Leu Gly Ile

770 775 780

Phe Leu Ala Tyr Glu Thr Lys Ser Val Ser Thr Glu Lys Ile Asn Asp
785 790 795 800

His Arg Ala Val Gly Met Ala Ile Tyr Asn Val Ala Val Leu Cys Leu
805 810 815

Ile Thr Ala Pro Val Thr Met Ile Leu Ser Ser Gln Gln Asp Ala Ala
820 825 830

Phe Ala Phe Ala Ser Leu Ala Ile Val Phe Ser Ser Tyr Ile Thr Leu
835 840 845

Val Val Leu Phe Val Pro Lys Met Arg Arg Leu Ile Thr Arg Gly Glu
850 855 860

Trp Gln Ser Glu Thr Gln Asp Thr Met Lys Thr Gly Ser Ser Thr Asn
865 870 875 880

Asn Asn Glu Glu Glu Lys Ser Arg Leu Leu Glu Lys Glu Asn Arg Glu
885 890 895

Leu Glu Lys Ile Ile Ala Glu Lys Glu Glu Arg Val Ser Glu Leu Arg
900 905 910

His Gln Leu Gln Ser Arg Gln Gln Leu Arg Ser Arg Arg His Pro Pro
915 920 925

Thr Pro Pro Asp Pro Ser Gly Gly Leu Pro Arg Gly Pro Ser Glu Pro
930 935 940

Pro Asp Arg Leu Ser Cys Asp Gly Ser Arg Val His Leu Leu Tyr Lys
945 950 955 960

(2) INFORMATION FOR SEQ ID NO: 46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2538 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Rattus norwegicus

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..2532

(x) PUBLICATION INFORMATION:
(A) AUTHORS: Kaupmann, K
Huggel, K
Heid, J
Flor, P
Bischoff, M
Mickel, S
McMaster, G
Angst, C
Bittiger, H
Froestl, W
(B) TITLE: Expression cloning of GABA-B receptors uncovers similarity to metabotropic glutamate receptors
(C) JOURNAL: Nature
(D) VOLUME: 386
(F) PAGES: 239-246
(G) DATE: 20 March-1997

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

```
ATG GGC CCG GGG GGA CCC TGT ACC CCA GTG GGG TGG CCG CTG CCT CTT      48
Met Gly Pro Gly Gly Pro Cys Thr Pro Val Gly Trp Pro Leu Pro Leu
 1               5                  10                  15

CTG CTG GTG ATG GCG GCT GGG GTG GCT CCG GTG TGG GCC TCT CAC TCC      96
Leu Leu Val Met Ala Ala Gly Val Ala Pro Val Trp Ala Ser His Ser
            20                  25                  30
```

```
CCT CAT CTC CCG CGG CCT CAC CCG AGG GTC CCC CCG CAC CCC TCC TCA      144
Pro His Leu Pro Arg Pro His Pro Arg Val Pro Pro His Pro Ser Ser
          35                  40                  45

GAA CGG CGT GCA GTA TAC ATC GGG GCG CTG TTT CCC ATG AGC GGG GGC      192
Glu Arg Arg Ala Val Tyr Ile Gly Ala Leu Phe Pro Met Ser Gly Gly
     50                  55                  60

TGG CCG GGG GGC CAG GCC TGC CAG CCC GCG GTG GAG ATG GCG CTG GAG      240
Trp Pro Gly Gly Gln Ala Cys Gln Pro Ala Val Glu Met Ala Leu Glu
 65                  70                  75                  80

GAC GTT AAC AGC CGC AGA GAC ATC CTG CCG GAC TAC GAG CTC AAG CTT      288
Asp Val Asn Ser Arg Arg Asp Ile Leu Pro Asp Tyr Glu Leu Lys Leu
                 85                  90                  95

ATC CAC CAC GAC AGC AAG TGT GAC CCA GGG CAA GCC ACC AAG TAC TTG      336
Ile His His Asp Ser Lys Cys Asp Pro Gly Gln Ala Thr Lys Tyr Leu
            100                 105                 110

TAC GAA CTA CTC TAC AAT GAC CCC ATC AAG ATC ATT CTC ATG CCT GGC      384
Tyr Glu Leu Leu Tyr Asn Asp Pro Ile Lys Ile Ile Leu Met Pro Gly
        115                 120                 125

TGT AGT TCT GTC TCC ACA CTT GTA GCT GAG GCT GCC CGG ATG TGG AAC      432
Cys Ser Ser Val Ser Thr Leu Val Ala Glu Ala Ala Arg Met Trp Asn
    130                 135                 140

CTT ATT GTG CTC TCA TAT GGC TCC AGT TCA CCA GCC TTG TCA AAC CGA      480
Leu Ile Val Leu Ser Tyr Gly Ser Ser Ser Pro Ala Leu Ser Asn Arg
145                 150                 155                 160

CAG CGG TTT CCC ACG TTC TTC CGG ACG CAT CCA TCC GCC ACA CTC CAC      528
Gln Arg Phe Pro Thr Phe Phe Arg Thr His Pro Ser Ala Thr Leu His
                165                 170                 175

AAT CCC ACC CGG GTG AAA CTC TTC GAA AAG TGG GGC TGG AAG AAG ATC      576
Asn Pro Thr Arg Val Lys Leu Phe Glu Lys Trp Gly Trp Lys Lys Ile
            180                 185                 190

GCT ACC ATC CAA CAG ACC ACC GAG GTC TTC ACC TCA ACG CTG GAT GAC      624
Ala Thr Ile Gln Gln Thr Thr Glu Val Phe Thr Ser Thr Leu Asp Asp
        195                 200                 205
```

```
CTG GAG GAG CGA GTG AAA GAG GCT GGG ATC GAG ATC ACT TTC CGA CAG      672
Leu Glu Glu Arg Val Lys Glu Ala Gly Ile Glu Ile Thr Phe Arg Gln
    210                 215                 220

AGT TTC TTC TCG GAT CCA GCT GTG CCT GTT AAA AAC CTG AAG CGT CAA      720
Ser Phe Phe Ser Asp Pro Ala Val Pro Val Lys Asn Leu Lys Arg Gln
225                 230                 235                 240

GAT GCT CGA ATC ATC GTG GGA CTT TTC TAT GAG ACG GAA GCC CGG AAA      768
Asp Ala Arg Ile Ile Val Gly Leu Phe Tyr Glu Thr Glu Ala Arg Lys
                245                 250                 255

GTT TTT TGT GAG GTC TAT AAG GAA AGG CTC TTT GGG AAG AAG TAC GTC      816
Val Phe Cys Glu Val Tyr Lys Glu Arg Leu Phe Gly Lys Lys Tyr Val
            260                 265                 270

TGG TTC CTC ATC GGG TGG TAT GCT GAC AAC TGG TTC AAG ACC TAT GAC      864
Trp Phe Leu Ile Gly Trp Tyr Ala Asp Asn Trp Phe Lys Thr Tyr Asp
        275                 280                 285

CCG TCA ATC AAT TGT ACA GTG GAA GAA ATG ACC GAG GCG GTG GAG GGC      912
Pro Ser Ile Asn Cys Thr Val Glu Glu Met Thr Glu Ala Val Glu Gly
    290                 295                 300

CAC ATC ACC ACG GAG ATT GTC ATG CTG AAC CCT GCC AAC ACC CGA AGC      960
His Ile Thr Thr Glu Ile Val Met Leu Asn Pro Ala Asn Thr Arg Ser
305                 310                 315                 320

ATT TCC AAC ATG ACG TCA CAG GAA TTT GTG GAG AAA CTA ACC AAG CGG     1008
Ile Ser Asn Met Thr Ser Gln Glu Phe Val Glu Lys Leu Thr Lys Arg
                325                 330                 335

CTG AAA AGA CAC CCC GAG GAG ACT GGA GGC TTC CAG GAG GCA CCA CTG     1056
Leu Lys Arg His Pro Glu Glu Thr Gly Gly Phe Gln Glu Ala Pro Leu
            340                 345                 350

GCC TAT GAT GCT ATC TGG GCC TTG GCT TTG GCC TTG AAC AAG ACG TCT     1104
Ala Tyr Asp Ala Ile Trp Ala Leu Ala Leu Ala Leu Asn Lys Thr Ser
        355                 360                 365

GGA GGA GGT GGT CGT TCC GGC GTG CGC CTG GAG GAC TTT AAC TAC AAC     1152
Gly Gly Gly Gly Arg Ser Gly Val Arg Leu Glu Asp Phe Asn Tyr Asn
```

```
    370                 375                 380

    AAC CAG ACC ATT ACA GAC CAG ATC TAC CGG GCC ATG AAC TCC TCC TCC      1200
    Asn Gln Thr Ile Thr Asp Gln Ile Tyr Arg Ala Met Asn Ser Ser Ser
5   385                 390                 395                 400

    TTT GAG GGC GTT TCT GGC CAT GTG GTC TTT GAT GCC AGC GGC TCC CGG      1248
    Phe Glu Gly Val Ser Gly His Val Val Phe Asp Ala Ser Gly Ser Arg
                    405                 410                 415

10

    ATG GCA TGG ACA CTT ATC GAG CAG CTA CAG GGC GGC AGC TAC AAG AAG      1296
    Met Ala Trp Thr Leu Ile Glu Gln Leu Gln Gly Gly Ser Tyr Lys Lys
                420                 425                 430

15  ATC GGC TAC TAC GAC AGC ACC AAG GAT GAT CTT TCC TGG TCC AAA ACG      1344
    Ile Gly Tyr Tyr Asp Ser Thr Lys Asp Asp Leu Ser Trp Ser Lys Thr
            435                 440                 445

    GAC AAG TGG ATT GGA GGG TCT CCC CCA GCT GAC CAG ACC TTG GTC ATC      1392
20  Asp Lys Trp Ile Gly Gly Ser Pro Pro Ala Asp Gln Thr Leu Val Ile
        450                 455                 460

    AAG ACA TTC CGT TTC CTG TCT CAG AAA CTC TTT ATC TCC GTC TCA GTT      1440
    Lys Thr Phe Arg Phe Leu Ser Gln Lys Leu Phe Ile Ser Val Ser Val
25  465                 470                 475                 480

    CTC TCC AGC CTG GGC ATT GTT CTT GCT GTT GTC TGT CTG TCC TTT AAC      1488
    Leu Ser Ser Leu Gly Ile Val Leu Ala Val Val Cys Leu Ser Phe Asn
                    485                 490                 495

30

    ATC TAC AAC TCC CAC GTT CGT TAT ATC CAG AAC TCC CAG CCC AAC CTG      1536
    Ile Tyr Asn Ser His Val Arg Tyr Ile Gln Asn Ser Gln Pro Asn Leu
                500                 505                 510

35  AAC AAT CTG ACT GCT GTG GGC TGC TCA CTG GCA CTG GCT GCT GTC TTC      1584
    Asn Asn Leu Thr Ala Val Gly Cys Ser Leu Ala Leu Ala Ala Val Phe
            515                 520                 525

    CCT CTC GGG CTG GAT GGT TAC CAC ATA GGG AGA AGC CAG TTC CCG TTT      1632
40  Pro Leu Gly Leu Asp Gly Tyr His Ile Gly Arg Ser Gln Phe Pro Phe
        530                 535                 540

    GTC TGC CAG GCC CGC CTT TGG CTC TTG GGC TTG GGC TTT AGT CTG GGC      1680
```

```
Val Cys Gln Ala Arg Leu Trp Leu Leu Gly Leu Gly Phe Ser Leu Gly
545                 550                 555                 560

TAT GGC TCT ATG TTC ACC AAG ATC TGG TGG GTC CAC ACA GTC TTC ACG      1728
Tyr Gly Ser Met Phe Thr Lys Ile Trp Trp Val His Thr Val Phe Thr
                565                 570                 575

AAG AAG GAG GAG AAG AAG GAG TGG AGG AAG ACC CTA GAG CCC TGG AAA      1776
Lys Lys Glu Glu Lys Lys Glu Trp Arg Lys Thr Leu Glu Pro Trp Lys
            580                 585                 590

CTC TAT GCC ACT GTG GGC CTG CTG GTG GGC ATG GAT GTC CTG ACT CTT      1824
Leu Tyr Ala Thr Val Gly Leu Leu Val Gly Met Asp Val Leu Thr Leu
        595                 600                 605

GCC ATC TGG CAG ATT GTG GAC CCC TTG CAC CGA ACC ATT GAG ACT TTT      1872
Ala Ile Trp Gln Ile Val Asp Pro Leu His Arg Thr Ile Glu Thr Phe
    610                 615                 620

GCC AAG GAG GAA CCA AAG GAA GAC ATC GAT GTC TCC ATT CTG CCC CAG      1920
Ala Lys Glu Glu Pro Lys Glu Asp Ile Asp Val Ser Ile Leu Pro Gln
625                 630                 635                 640

TTG GAG CAC TGC AGC TCC AAG AAG ATG AAT ACG TGG CTT GGC ATT TTC      1968
Leu Glu His Cys Ser Ser Lys Lys Met Asn Thr Trp Leu Gly Ile Phe
                645                 650                 655

TAT GGT TAC AAG GGG CTG CTG CTG CTG CTG GGA ATC TTT CTT GCT TAC      2016
Tyr Gly Tyr Lys Gly Leu Leu Leu Leu Leu Gly Ile Phe Leu Ala Tyr
            660                 665                 670

GAA ACC AAG AGC GTG TCC ACT GAA AAG ATC AAT GAC CAC AGG GCC GTG      2064
Glu Thr Lys Ser Val Ser Thr Glu Lys Ile Asn Asp His Arg Ala Val
        675                 680                 685

GGC ATG GCT ATC TAC AAT GTC GCG GTC CTG TGT CTC ATC ACT GCT CCT      2112
Gly Met Ala Ile Tyr Asn Val Ala Val Leu Cys Leu Ile Thr Ala Pro
    690                 695                 700

GTG ACC ATG ATC CTT TCC AGT CAG CAG GAC GCA GCC TTT GCC TTT GCC      2160
Val Thr Met Ile Leu Ser Ser Gln Gln Asp Ala Ala Phe Ala Phe Ala
705                 710                 715                 720
```

```
TCT CTG GCC ATC GTG TTC TCT TCC TAC ATC ACT CTG GTT GTG CTC TTT      2208
Ser Leu Ala Ile Val Phe Ser Ser Tyr Ile Thr Leu Val Val Leu Phe
                725                 730                 735

GTG CCC AAG ATG CGC AGG CTG ATC ACC CGA GGG GAA TGG CAG TCT GAA      2256
Val Pro Lys Met Arg Arg Leu Ile Thr Arg Gly Glu Trp Gln Ser Glu
            740                 745                 750

ACG CAG GAC ACC ATG AAA ACA GGA TCA TCC ACC AAC AAC AAC GAG GAA      2304
Thr Gln Asp Thr Met Lys Thr Gly Ser Ser Thr Asn Asn Asn Glu Glu
        755                 760                 765

GAG AAG TCC CGA CTG TTG GAG AAG GAA AAC CGA GAA CTG GAA AAG ATC      2352
Glu Lys Ser Arg Leu Leu Glu Lys Glu Asn Arg Glu Leu Glu Lys Ile
    770                 775                 780

ATC GCT GAG AAA GAG GAG CGC GTC TCT GAA CTG CGC CAT CAG CTC CAG      2400
Ile Ala Glu Lys Glu Glu Arg Val Ser Glu Leu Arg His Gln Leu Gln
785                 790                 795                 800

TCT CGG CAG CAA CTC CGC TCA CGG CGC CAC CCC CCA ACA CCC CCA GAT      2448
Ser Arg Gln Gln Leu Arg Ser Arg Arg His Pro Pro Thr Pro Pro Asp
                805                 810                 815

CCC TCT GGG GGC CTT CCC AGG GGA CCC TCT GAG CCC CCT GAC CGG CTT      2496
Pro Ser Gly Gly Leu Pro Arg Gly Pro Ser Glu Pro Pro Asp Arg Leu
            820                 825                 830

AGC TGT GAT GGG AGT CGA GTA CAT TTG CTT TAC AAG TGA                  2535
Ser Cys Asp Gly Ser Arg Val His Leu Leu Tyr Lys  *
        835                 840
```

(2) INFORMATION FOR SEQ ID NO: 47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 844 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

```
Met Gly Pro Gly Gly Pro Cys Thr Pro Val Gly Trp Pro Leu Pro Leu
 1               5                  10                  15

Leu Leu Val Met Ala Ala Gly Val Ala Pro Val Trp Ala Ser His Ser
            20                  25                  30

Pro His Leu Pro Arg Pro His Pro Arg Val Pro Pro His Pro Ser Ser
        35                  40                  45

Glu Arg Arg Ala Val Tyr Ile Gly Ala Leu Phe Pro Met Ser Gly Gly
    50                  55                  60

Trp Pro Gly Gly Gln Ala Cys Gln Pro Ala Val Glu Met Ala Leu Glu
65                  70                  75                  80

Asp Val Asn Ser Arg Arg Asp Ile Leu Pro Asp Tyr Glu Leu Lys Leu
                85                  90                  95

Ile His His Asp Ser Lys Cys Asp Pro Gly Gln Ala Thr Lys Tyr Leu
            100                 105                 110

Tyr Glu Leu Leu Tyr Asn Asp Pro Ile Lys Ile Ile Leu Met Pro Gly
        115                 120                 125

Cys Ser Ser Val Ser Thr Leu Val Ala Glu Ala Ala Arg Met Trp Asn
    130                 135                 140

Leu Ile Val Leu Ser Tyr Gly Ser Ser Ser Pro Ala Leu Ser Asn Arg
145                 150                 155                 160

Gln Arg Phe Pro Thr Phe Phe Arg Thr His Pro Ser Ala Thr Leu His
                165                 170                 175

Asn Pro Thr Arg Val Lys Leu Phe Glu Lys Trp Gly Trp Lys Lys Ile
            180                 185                 190

Ala Thr Ile Gln Gln Thr Thr Glu Val Phe Thr Ser Thr Leu Asp Asp
        195                 200                 205

Leu Glu Glu Arg Val Lys Glu Ala Gly Ile Glu Ile Thr Phe Arg Gln
    210                 215                 220
```

```
Ser Phe Phe Ser Asp Pro Ala Val Pro Val Lys Asn Leu Lys Arg Gln
225                 230                 235                 240

Asp Ala Arg Ile Ile Val Gly Leu Phe Tyr Glu Thr Glu Ala Arg Lys
                245                 250                 255

Val Phe Cys Glu Val Tyr Lys Glu Arg Leu Phe Gly Lys Lys Tyr Val
            260                 265                 270

Trp Phe Leu Ile Gly Trp Tyr Ala Asp Asn Trp Phe Lys Thr Tyr Asp
        275                 280                 285

Pro Ser Ile Asn Cys Thr Val Glu Glu Met Thr Glu Ala Val Glu Gly
    290                 295                 300

His Ile Thr Thr Glu Ile Val Met Leu Asn Pro Ala Asn Thr Arg Ser
305                 310                 315                 320

Ile Ser Asn Met Thr Ser Gln Glu Phe Val Glu Lys Leu Thr Lys Arg
                325                 330                 335

Leu Lys Arg His Pro Glu Glu Thr Gly Gly Phe Gln Glu Ala Pro Leu
            340                 345                 350

Ala Tyr Asp Ala Ile Trp Ala Leu Ala Leu Ala Leu Asn Lys Thr Ser
        355                 360                 365

Gly Gly Gly Gly Arg Ser Gly Val Arg Leu Glu Asp Phe Asn Tyr Asn
    370                 375                 380

Asn Gln Thr Ile Thr Asp Gln Ile Tyr Arg Ala Met Asn Ser Ser Ser
385                 390                 395                 400

Phe Glu Gly Val Ser Gly His Val Val Phe Asp Ala Ser Gly Ser Arg
                405                 410                 415

Met Ala Trp Thr Leu Ile Glu Gln Leu Gln Gly Gly Ser Tyr Lys Lys
            420                 425                 430

Ile Gly Tyr Tyr Asp Ser Thr Lys Asp Asp Leu Ser Trp Ser Lys Thr
        435                 440                 445

Asp Lys Trp Ile Gly Gly Ser Pro Pro Ala Asp Gln Thr Leu Val Ile
```

450 455 460

Lys Thr Phe Arg Phe Leu Ser Gln Lys Leu Phe Ile Ser Val Ser Val
465 470 475 480

Leu Ser Ser Leu Gly Ile Val Leu Ala Val Val Cys Leu Ser Phe Asn
485 490 495

Ile Tyr Asn Ser His Val Arg Tyr Ile Gln Asn Ser Gln Pro Asn Leu
500 505 510

Asn Asn Leu Thr Ala Val Gly Cys Ser Leu Ala Leu Ala Ala Val Phe
515 520 525

Pro Leu Gly Leu Asp Gly Tyr His Ile Gly Arg Ser Gln Phe Pro Phe
530 535 540

Val Cys Gln Ala Arg Leu Trp Leu Leu Gly Leu Gly Phe Ser Leu Gly
545 550 555 560

Tyr Gly Ser Met Phe Thr Lys Ile Trp Trp Val His Thr Val Phe Thr
565 570 575

Lys Lys Glu Glu Lys Lys Glu Trp Arg Lys Thr Leu Glu Pro Trp Lys
580 585 590

Leu Tyr Ala Thr Val Gly Leu Leu Val Gly Met Asp Val Leu Thr Leu
595 600 605

Ala Ile Trp Gln Ile Val Asp Pro Leu His Arg Thr Ile Glu Thr Phe
610 615 620

Ala Lys Glu Glu Pro Lys Glu Asp Ile Asp Val Ser Ile Leu Pro Gln
625 630 635 640

Leu Glu His Cys Ser Ser Lys Lys Met Asn Thr Trp Leu Gly Ile Phe
645 650 655

Tyr Gly Tyr Lys Gly Leu Leu Leu Leu Leu Gly Ile Phe Leu Ala Tyr
660 665 670

Glu Thr Lys Ser Val Ser Thr Glu Lys Ile Asn Asp His Arg Ala Val
675 680 685

```
Gly Met Ala Ile Tyr Asn Val Ala Val Leu Cys Leu Ile Thr Ala Pro
    690                 695                 700

Val Thr Met Ile Leu Ser Ser Gln Gln Asp Ala Ala Phe Ala Phe Ala
705                 710                 715                 720

Ser Leu Ala Ile Val Phe Ser Ser Tyr Ile Thr Leu Val Val Leu Phe
                725                 730                 735

Val Pro Lys Met Arg Arg Leu Ile Thr Arg Gly Glu Trp Gln Ser Glu
            740                 745                 750

Thr Gln Asp Thr Met Lys Thr Gly Ser Ser Thr Asn Asn Asn Glu Glu
        755                 760                 765

Glu Lys Ser Arg Leu Leu Glu Lys Glu Asn Arg Glu Leu Glu Lys Ile
    770                 775                 780

Ile Ala Glu Lys Glu Glu Arg Val Ser Glu Leu Arg His Gln Leu Gln
785                 790                 795                 800

Ser Arg Gln Gln Leu Arg Ser Arg Arg His Pro Pro Thr Pro Pro Asp
                805                 810                 815

Pro Ser Gly Gly Leu Pro Arg Gly Pro Ser Glu Pro Pro Asp Arg Leu
            820                 825                 830

Ser Cys Asp Gly Ser Arg Val His Leu Leu Tyr Lys
        835                 840
```

(2) INFORMATION FOR SEQ ID NO: 48:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2886 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..2886

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

```
ATG TTG CTG CTG CTG CTA CTG GCG CCA CTC TTC CTC CGC CCC CCG GGC          48
Met Leu Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
 1               5                  10                  15

GCG GGC GGG GCG CAG ACC CCC AAC GCC ACC TCA GAA GGT TGC CAG ATC          96
Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
            20                  25                  30

ATA CAC CCG CCC TGG GAA GGG GGC ATC AGG TAC CGG GGC CTG ACT CGG         144
Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
        35                  40                  45

GAC CAG GTG AAG GCT ATC AAC TTC CTG CCA GTG GAC TAT GAG ATT GAG         192
Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu
    50                  55                  60

TAT GTG TGC CGG GGG GAG CGC GAG GTG GTG GGG CCC AAG GTC CGC AAG         240
Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys
65                  70                  75                  80

TGC CTG GCC AAC GGC TCC TGG ACA GAT ATG GAC ACA CCC AGC CGC TGT         288
Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys
                85                  90                  95

GTC CGA ATC TGC TCC AAG TCT TAT TTG ACC CTG GAA AAT GGG AAG GTT         336
Val Arg Ile Cys Ser Lys Ser Tyr Leu Thr Leu Glu Asn Gly Lys Val
            100                 105                 110

TTC CTG ACG GGT GGG GAC CTC CCA GCT CTG GAC GGA GCC CGG GTG GAT         384
Phe Leu Thr Gly Gly Asp Leu Pro Ala Leu Asp Gly Ala Arg Val Asp
        115                 120                 125
```

```
TTC CGG TGT GAC CCC GAC TTC CAT CTG GTG GGC AGC TCC CGG AGC ATC      432
Phe Arg Cys Asp Pro Asp Phe His Leu Val Gly Ser Ser Arg Ser Ile
    130                 135                 140

TGT AGT CAG GGC CAG TGG AGC ACC CCC AAG CCC CAC TGC CAG GTG AAT      480
Cys Ser Gln Gly Gln Trp Ser Thr Pro Lys Pro His Cys Gln Val Asn
145                 150                 155                 160

CGA ACG CCA CAC TCA GAA CGG CGC GCA GTG TAC ATC GGG GCA CTG TTT      528
Arg Thr Pro His Ser Glu Arg Arg Ala Val Tyr Ile Gly Ala Leu Phe
                165                 170                 175

CCC ATG AGC GGG GGC TGG CCA GGG GGC CAG GCC TGC CAG CCC GCG GTG      576
Pro Met Ser Gly Gly Trp Pro Gly Gly Gln Ala Cys Gln Pro Ala Val
            180                 185                 190

GAG ATG GCG CTG GAG GAC GTG AAT AGC CGC AGG GAC ATC CTG CCG GAC      624
Glu Met Ala Leu Glu Asp Val Asn Ser Arg Arg Asp Ile Leu Pro Asp
        195                 200                 205

TAT GAG CTC AAG CTC ATC CAC CAC GAC AGC AAG TGT GAT CCA GGC CAA      672
Tyr Glu Leu Lys Leu Ile His His Asp Ser Lys Cys Asp Pro Gly Gln
    210                 215                 220

GCC ACC AAG TAC CTA TAT GAG CTG CTC TAC AAC GAC CCT ATC AAG ATC      720
Ala Thr Lys Tyr Leu Tyr Glu Leu Leu Tyr Asn Asp Pro Ile Lys Ile
225                 230                 235                 240

ATC CTT ATG CCT GGC TGC AGC TCT GTC TCC ACG CTG GTG GCT GAG GCT      768
Ile Leu Met Pro Gly Cys Ser Ser Val Ser Thr Leu Val Ala Glu Ala
                245                 250                 255

GCT AGG ATG TGG AAC CTC ATT GTG CTT TCC TAT GGC TCC AGC TCA CCA      816
Ala Arg Met Trp Asn Leu Ile Val Leu Ser Tyr Gly Ser Ser Ser Pro
            260                 265                 270

GCC CTG TCA AAC CGG CAG CGT TTC CCC ACT TTC TTC CGA ACG CAC CCA      864
Ala Leu Ser Asn Arg Gln Arg Phe Pro Thr Phe Phe Arg Thr His Pro
        275                 280                 285

TCA GCC ACA CTC CAC AAC CCT ACC CGC GTG AAA CTC TTT GAA AAG TGG      912
Ser Ala Thr Leu His Asn Pro Thr Arg Val Lys Leu Phe Glu Lys Trp
```

290 295 300

GGC TGG AAG AAG ATT GCT ACC ATC CAG CAG ACC ACT GAG GTC TTC ACT 960
Gly Trp Lys Lys Ile Ala Thr Ile Gln Gln Thr Thr Glu Val Phe Thr
305 310 315 320

TCG ACT CTG GAC GAC CTG GAG GAA CGA GTG AAG GAG GCT GGA ATT GAG 1008
Ser Thr Leu Asp Asp Leu Glu Glu Arg Val Lys Glu Ala Gly Ile Glu
325 330 335

ATT ACT TTC CGC CAG AGT TTC TTC TCA GAT CCA GCT GTG CCC GTC AAA 1056
Ile Thr Phe Arg Gln Ser Phe Phe Ser Asp Pro Ala Val Pro Val Lys
340 345 350

AAC CTG AAG CGC CAG GAT GCC CGA ATC ATC GTG GGA CTT TTC TAT GAG 1104
Asn Leu Lys Arg Gln Asp Ala Arg Ile Ile Val Gly Leu Phe Tyr Glu
355 360 365

ACT GAA GCC CGG AAA GTT TTT TGT GAG GTG TAC AAG GAG CGT CTC TTT 1152
Thr Glu Ala Arg Lys Val Phe Cys Glu Val Tyr Lys Glu Arg Leu Phe
370 375 380

GGG AAG AAG TAC GTC TGG TTC CTC ATT GGG TGG TAT GCT GAC AAT TGG 1200
Gly Lys Lys Tyr Val Trp Phe Leu Ile Gly Trp Tyr Ala Asp Asn Trp
385 390 395 400

TTC AAG ATC TAC GAC CCT TCT ATC AAC TGC ACA GTG GAT GAG ATG ACT 1248
Phe Lys Ile Tyr Asp Pro Ser Ile Asn Cys Thr Val Asp Glu Met Thr
405 410 415

GAG GCG GTG GAG GGC CAC ATC ACA ACT GAG ATT GTC ATG CTG AAT CCT 1296
Glu Ala Val Glu Gly His Ile Thr Thr Glu Ile Val Met Leu Asn Pro
420 425 430

GCC AAT ACC CGC AGC ATT TCC AAC ATG ACA TCC CAG GAA TTT GTG GAG 1344
Ala Asn Thr Arg Ser Ile Ser Asn Met Thr Ser Gln Glu Phe Val Glu
435 440 445

AAA CTA ACC AAG CGA CTG AAA AGA CAC CCT GAG GAG ACA GGA GGC TTC 1392
Lys Leu Thr Lys Arg Leu Lys Arg His Pro Glu Glu Thr Gly Gly Phe
450 455 460

CAG GAG GCA CCG CTG GCC TAT GAT GCC ATC TGG GCC TTG GCA CTG GCC 1440

```
Gln Glu Ala Pro Leu Ala Tyr Asp Ala Ile Trp Ala Leu Ala Leu Ala
465                 470                 475                 480

CTG AAC AAG ACA TCT GGA GGA GGC GGC CGT TCT GGT GTG CGC CTG GAG      1488
Leu Asn Lys Thr Ser Gly Gly Gly Gly Arg Ser Gly Val Arg Leu Glu
                485                 490                 495

GAC TTC AAC TAC AAC AAC CAG ACC ATT ACC GAC CAA ATC TAC CGG GCA      1536
Asp Phe Asn Tyr Asn Asn Gln Thr Ile Thr Asp Gln Ile Tyr Arg Ala
            500                 505                 510

ATG AAC TCT TCG TCC TTT GAG GGT GTC TCT GGC CAT GTG GTG TTT GAT      1584
Met Asn Ser Ser Ser Phe Glu Gly Val Ser Gly His Val Val Phe Asp
        515                 520                 525

GCC AGC GGC TCT CGG ATG GCA TGG ACG CTT ATC GAG CAG CTT CAG GGT      1632
Ala Ser Gly Ser Arg Met Ala Trp Thr Leu Ile Glu Gln Leu Gln Gly
    530                 535                 540

GGC AGC TAC AAG AAG ATT GGC TAC TAT GAC AGC ACC AAG GAT GAT CTT      1680
Gly Ser Tyr Lys Lys Ile Gly Tyr Tyr Asp Ser Thr Lys Asp Asp Leu
545                 550                 555                 560

TCC TGG TCC AAA ACA GAT AAA TGG ATT GGA GGG TCC CCC CCA GCT GAC      1728
Ser Trp Ser Lys Thr Asp Lys Trp Ile Gly Gly Ser Pro Pro Ala Asp
                565                 570                 575

CAG ACC CTG GTC ATC AAG ACA TTC CGC TTC CTG TCA CAG AAA CTC TTT      1776
Gln Thr Leu Val Ile Lys Thr Phe Arg Phe Leu Ser Gln Lys Leu Phe
            580                 585                 590

ATC TCC GTC TCA GTT CTC TCC AGC CTG GGC ATT GTC CTA GCT GTT GTC      1824
Ile Ser Val Ser Val Leu Ser Ser Leu Gly Ile Val Leu Ala Val Val
        595                 600                 605

TGT CTG TCC TTT AAC ATC TAC AAC TCA CAT GTC CGT TAT ATC CAG AAC      1872
Cys Leu Ser Phe Asn Ile Tyr Asn Ser His Val Arg Tyr Ile Gln Asn
    610                 615                 620

TCA CAG CCC AAC CTG AAC AAC CTG ACT GCT GTG GGC TGC TCA CTG GCT      1920
Ser Gln Pro Asn Leu Asn Asn Leu Thr Ala Val Gly Cys Ser Leu Ala
625                 630                 635                 640
```

```
TTA GCT GCT GTC TTC CCC CTG GGG CTC GAT GGT TAC CAC ATT GGG AGG      1968
Leu Ala Ala Val Phe Pro Leu Gly Leu Asp Gly Tyr His Ile Gly Arg
                645                 650                 655

AAC CAG TTT CCT TTC GTC TGC CAG GCC CGC CTC TGG CTC CTG GGC CTG      2016
Asn Gln Phe Pro Phe Val Cys Gln Ala Arg Leu Trp Leu Leu Gly Leu
            660                 665                 670

GGC TTT AGT CTG GGC TAC GGT TCC ATG TTC ACC AAG ATT TGG TGG GTC      2064
Gly Phe Ser Leu Gly Tyr Gly Ser Met Phe Thr Lys Ile Trp Trp Val
        675                 680                 685

CAC ACG GTC TTC ACA AAG AAG GAA GAA AAG AAG GAG TGG AGG AAG ACT      2112
His Thr Val Phe Thr Lys Lys Glu Glu Lys Lys Glu Trp Arg Lys Thr
    690                 695                 700

CTG GAA CCC TGG AAG CTG TAT GCC ACA GTG GGC CTG CTG GTG GGC ATG      2160
Leu Glu Pro Trp Lys Leu Tyr Ala Thr Val Gly Leu Leu Val Gly Met
705                 710                 715                 720

GAT GTC CTC ACT CTC GCC ATC TGG CAG ATC GTG GAC CCT CTG CAC CGG      2208
Asp Val Leu Thr Leu Ala Ile Trp Gln Ile Val Asp Pro Leu His Arg
                725                 730                 735

ACC ATT GAG ACA TTT GCC AAG GAG GAA CCT AAG GAA GAT ATT GAC GTC      2256
Thr Ile Glu Thr Phe Ala Lys Glu Glu Pro Lys Glu Asp Ile Asp Val
            740                 745                 750

TCT ATT CTG CCC CAG CTG GAG CAT TGC AGC TCC AGG AAG ATG AAT ACA      2304
Ser Ile Leu Pro Gln Leu Glu His Cys Ser Ser Arg Lys Met Asn Thr
        755                 760                 765

TGG CTT GGC ATT TTC TAT GGT TAC AAG GGG CTG CTG CTG CTG CTG GGA      2352
Trp Leu Gly Ile Phe Tyr Gly Tyr Lys Gly Leu Leu Leu Leu Leu Gly
    770                 775                 780

ATC TTC CTT GCT TAT GAG ACC AAG AGT GTG TCC ACT GAG AAG ATC AAT      2400
Ile Phe Leu Ala Tyr Glu Thr Lys Ser Val Ser Thr Glu Lys Ile Asn
785                 790                 795                 800

GAT CAC CGG GCT GTG GGC ATG GCT ATC TAC AAT GTG GCA GTC CTG TGC      2448
Asp His Arg Ala Val Gly Met Ala Ile Tyr Asn Val Ala Val Leu Cys
                805                 810                 815
```

```
CTC ATC ACT GCT CCT GTC ACC ATG ATT CTG TCC AGC CAG CAG GAT GCA      2496
Leu Ile Thr Ala Pro Val Thr Met Ile Leu Ser Ser Gln Gln Asp Ala
            820                 825                 830

GCC TTT GCC TTT GCC TCT CTT GCC ATA GTT TTC TCC TCC TAT ATC ACT      2544
Ala Phe Ala Phe Ala Ser Leu Ala Ile Val Phe Ser Ser Tyr Ile Thr
        835                 840                 845

CTT GTT GTG CTC TTT GTG CCC AAG ATG CGC AGG CTG ATC ACC CGA GGG      2592
Leu Val Val Leu Phe Val Pro Lys Met Arg Arg Leu Ile Thr Arg Gly
    850                 855                 860

GAA TGG CAG TCG GAG GCG CAG GAC ACC ATG AAG ACA GGG TCA TCG ACC      2640
Glu Trp Gln Ser Glu Ala Gln Asp Thr Met Lys Thr Gly Ser Ser Thr
865                 870                 875                 880

AAC AAC AAC GAG GAG GAG AAG TCC CGG CTG TTG GAG AAG GAG AAC CGT      2688
Asn Asn Asn Glu Glu Glu Lys Ser Arg Leu Leu Glu Lys Glu Asn Arg
                885                 890                 895

GAA CTG GAA AAG ATC ATT GCT GAG AAA GAG GAG CGT GTC TCT GAA CTG      2736
Glu Leu Glu Lys Ile Ile Ala Glu Lys Glu Glu Arg Val Ser Glu Leu
            900                 905                 910

CGC CAT CAA CTC CAG TCT CGG CAG CAG CTC CGC TCC CGG CGC CAC CCA      2784
Arg His Gln Leu Gln Ser Arg Gln Gln Leu Arg Ser Arg Arg His Pro
        915                 920                 925

CCG ACA CCC CCA GAA CCC TCT GGG GGC CTG CCC AGG GGA CCC CCT GAG      2832
Pro Thr Pro Pro Glu Pro Ser Gly Gly Leu Pro Arg Gly Pro Pro Glu
    930                 935                 940

CCC CCC GAC CGG CTT AGC TGT GAT GGG AGT CGA GTG CAT TTG CTT TAT      2880
Pro Pro Asp Arg Leu Ser Cys Asp Gly Ser Arg Val His Leu Leu Tyr
945                 950                 955                 960

AAG TGA                                                              2886
Lys  *
```

(2) INFORMATION FOR SEQ ID NO: 49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 961 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

Met Leu Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
1 5 10 15

Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
20 25 30

Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
35 40 45

Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu
50 55 60

Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys
65 70 75 80

Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys
85 90 95

Val Arg Ile Cys Ser Lys Ser Tyr Leu Thr Leu Glu Asn Gly Lys Val
100 105 110

Phe Leu Thr Gly Gly Asp Leu Pro Ala Leu Asp Gly Ala Arg Val Asp
115 120 125

Phe Arg Cys Asp Pro Asp Phe His Leu Val Gly Ser Ser Arg Ser Ile
130 135 140

Cys Ser Gln Gly Gln Trp Ser Thr Pro Lys Pro His Cys Gln Val Asn
145 150 155 160

Arg Thr Pro His Ser Glu Arg Arg Ala Val Tyr Ile Gly Ala Leu Phe
165 170 175

```
Pro Met Ser Gly Gly Trp Pro Gly Gly Gln Ala Cys Gln Pro Ala Val
            180                 185                 190

Glu Met Ala Leu Glu Asp Val Asn Ser Arg Arg Asp Ile Leu Pro Asp
        195                 200                 205

Tyr Glu Leu Lys Leu Ile His His Asp Ser Lys Cys Asp Pro Gly Gln
    210                 215                 220

Ala Thr Lys Tyr Leu Tyr Glu Leu Leu Tyr Asn Asp Pro Ile Lys Ile
225                 230                 235                 240

Ile Leu Met Pro Gly Cys Ser Ser Val Ser Thr Leu Val Ala Glu Ala
                245                 250                 255

Ala Arg Met Trp Asn Leu Ile Val Leu Ser Tyr Gly Ser Ser Ser Pro
            260                 265                 270

Ala Leu Ser Asn Arg Gln Arg Phe Pro Thr Phe Phe Arg Thr His Pro
        275                 280                 285

Ser Ala Thr Leu His Asn Pro Thr Arg Val Lys Leu Phe Glu Lys Trp
    290                 295                 300

Gly Trp Lys Lys Ile Ala Thr Ile Gln Gln Thr Thr Glu Val Phe Thr
305                 310                 315                 320

Ser Thr Leu Asp Asp Leu Glu Glu Arg Val Lys Glu Ala Gly Ile Glu
                325                 330                 335

Ile Thr Phe Arg Gln Ser Phe Phe Ser Asp Pro Ala Val Pro Val Lys
            340                 345                 350

Asn Leu Lys Arg Gln Asp Ala Arg Ile Ile Val Gly Leu Phe Tyr Glu
        355                 360                 365

Thr Glu Ala Arg Lys Val Phe Cys Glu Val Tyr Lys Glu Arg Leu Phe
    370                 375                 380

Gly Lys Lys Tyr Val Trp Phe Leu Ile Gly Trp Tyr Ala Asp Asn Trp
385                 390                 395                 400

Phe Lys Ile Tyr Asp Pro Ser Ile Asn Cys Thr Val Asp Glu Met Thr
```

405 410 415

Glu Ala Val Glu Gly His Ile Thr Thr Glu Ile Val Met Leu Asn Pro
420 425 430

Ala Asn Thr Arg Ser Ile Ser Asn Met Thr Ser Gln Glu Phe Val Glu
435 440 445

Lys Leu Thr Lys Arg Leu Lys Arg His Pro Glu Glu Thr Gly Gly Phe
450 455 460

Gln Glu Ala Pro Leu Ala Tyr Asp Ala Ile Trp Ala Leu Ala Leu Ala
465 470 475 480

Leu Asn Lys Thr Ser Gly Gly Gly Gly Arg Ser Gly Val Arg Leu Glu
485 490 495

Asp Phe Asn Tyr Asn Asn Gln Thr Ile Thr Asp Gln Ile Tyr Arg Ala
500 505 510

Met Asn Ser Ser Ser Phe Glu Gly Val Ser Gly His Val Val Phe Asp
515 520 525

Ala Ser Gly Ser Arg Met Ala Trp Thr Leu Ile Glu Gln Leu Gln Gly
530 535 540

Gly Ser Tyr Lys Lys Ile Gly Tyr Tyr Asp Ser Thr Lys Asp Asp Leu
545 550 555 560

Ser Trp Ser Lys Thr Asp Lys Trp Ile Gly Gly Ser Pro Pro Ala Asp
565 570 575

Gln Thr Leu Val Ile Lys Thr Phe Arg Phe Leu Ser Gln Lys Leu Phe
580 585 590

Ile Ser Val Ser Val Leu Ser Ser Leu Gly Ile Val Leu Ala Val Val
595 600 605

Cys Leu Ser Phe Asn Ile Tyr Asn Ser His Val Arg Tyr Ile Gln Asn
610 615 620

Ser Gln Pro Asn Leu Asn Asn Leu Thr Ala Val Gly Cys Ser Leu Ala
625 630 635 640

```
Leu Ala Ala Val Phe Pro Leu Gly Leu Asp Gly Tyr His Ile Gly Arg
                645                 650                 655

Asn Gln Phe Pro Phe Val Cys Gln Ala Arg Leu Trp Leu Leu Gly Leu
            660                 665                 670

Gly Phe Ser Leu Gly Tyr Gly Ser Met Phe Thr Lys Ile Trp Trp Val
        675                 680                 685

His Thr Val Phe Thr Lys Lys Glu Glu Lys Lys Glu Trp Arg Lys Thr
    690                 695                 700

Leu Glu Pro Trp Lys Leu Tyr Ala Thr Val Gly Leu Leu Val Gly Met
705                 710                 715                 720

Asp Val Leu Thr Leu Ala Ile Trp Gln Ile Val Asp Pro Leu His Arg
                725                 730                 735

Thr Ile Glu Thr Phe Ala Lys Glu Glu Pro Lys Glu Asp Ile Asp Val
            740                 745                 750

Ser Ile Leu Pro Gln Leu Glu His Cys Ser Ser Arg Lys Met Asn Thr
        755                 760                 765

Trp Leu Gly Ile Phe Tyr Gly Tyr Lys Gly Leu Leu Leu Leu Leu Gly
    770                 775                 780

Ile Phe Leu Ala Tyr Glu Thr Lys Ser Val Ser Thr Glu Lys Ile Asn
785                 790                 795                 800

Asp His Arg Ala Val Gly Met Ala Ile Tyr Asn Val Ala Val Leu Cys
                805                 810                 815

Leu Ile Thr Ala Pro Val Thr Met Ile Leu Ser Ser Gln Gln Asp Ala
            820                 825                 830

Ala Phe Ala Phe Ala Ser Leu Ala Ile Val Phe Ser Ser Tyr Ile Thr
        835                 840                 845

Leu Val Val Leu Phe Val Pro Lys Met Arg Arg Leu Ile Thr Arg Gly
    850                 855                 860
```

Glu Trp Gln Ser Glu Ala Gln Asp Thr Met Lys Thr Gly Ser Ser Thr
865 870 875 880

Asn Asn Asn Glu Glu Glu Lys Ser Arg Leu Leu Glu Lys Glu Asn Arg
885 890 895

Glu Leu Glu Lys Ile Ile Ala Glu Lys Glu Glu Arg Val Ser Glu Leu
900 905 910

Arg His Gln Leu Gln Ser Arg Gln Gln Leu Arg Ser Arg Arg His Pro
915 920 925

Pro Thr Pro Pro Glu Pro Ser Gly Gly Leu Pro Arg Gly Pro Pro Glu
930 935 940

Pro Pro Asp Arg Leu Ser Cys Asp Gly Ser Arg Val His Leu Leu Tyr
945 950 955 960

Lys

(2) INFORMATION FOR SEQ ID NO: 50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2535 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to rRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..2535

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

```
ATG GGG CCC GGG GCC CCT TTT GCC CGG GTG GGG TGG CCA CTG CCG CTT   48
Met Gly Pro Gly Ala Pro Phe Ala Arg Val Gly Trp Pro Leu Pro Leu
 1               5                  10                  15

CTG GTT GTG ATG GCG GCA GGG GTG GCT CCG GTG TGG GCC TCC CAC TCC   96
Leu Val Val Met Ala Ala Gly Val Ala Pro Val Trp Ala Ser His Ser
            20                  25                  30

CCC CAT CTC CCG CGG CCT CAC TCG CGG GTC CCC CCG CAC CCC TCC TCA  144
Pro His Leu Pro Arg Pro His Ser Arg Val Pro Pro His Pro Ser Ser
        35                  40                  45

GAA CGG CGC GCA GTG TAC ATC GGG GCA CTG TTT CCC ATG AGC GGG GGC  192
Glu Arg Arg Ala Val Tyr Ile Gly Ala Leu Phe Pro Met Ser Gly Gly
    50                  55                  60

TGG CCA GGG GGC CAG GCC TGC CAG CCC GCG GTG GAG ATG GCG CTG GAG  240
Trp Pro Gly Gly Gln Ala Cys Gln Pro Ala Val Glu Met Ala Leu Glu
65                  70                  75                  80

GAC GTG AAT AGC CGC AGG GAC ATC CTG CCG GAC TAT GAG CTC AAG CTC  288
Asp Val Asn Ser Arg Arg Asp Ile Leu Pro Asp Tyr Glu Leu Lys Leu
                85                  90                  95

ATC CAC CAC GAC AGC AAG TGT GAT CCA GGC CAA GCC ACC AAG TAC CTA  336
Ile His His Asp Ser Lys Cys Asp Pro Gly Gln Ala Thr Lys Tyr Leu
            100                 105                 110

TAT GAG CTG CTC TAC AAC GAC CCT ATC AAG ATC ATC CTT ATG CCT GGC  384
Tyr Glu Leu Leu Tyr Asn Asp Pro Ile Lys Ile Ile Leu Met Pro Gly
        115                 120                 125

TGC AGC TCT GTC TCC ACG CTG GTG GCT GAG GCT GCT AGG ATG TGG AAC  432
Cys Ser Ser Val Ser Thr Leu Val Ala Glu Ala Ala Arg Met Trp Asn
    130                 135                 140

CTC ATT GTG CTT TCC TAT GGC TCC AGC TCA CCA GCC CTG TCA AAC CGG  480
Leu Ile Val Leu Ser Tyr Gly Ser Ser Ser Pro Ala Leu Ser Asn Arg
145                 150                 155                 160

CAG CGT TTC CCC ACT TTC TTC CGA ACG CAC CCA TCA GCC ACA CTC CAC  528
```

```
Gln Arg Phe Pro Thr Phe Phe Arg Thr His Pro Ser Ala Thr Leu His
                165                 170                 175

AAC CCT ACC CGC GTG AAA CTC TTT GAA AAG TGG GGC TGG AAG AAG ATT   576
Asn Pro Thr Arg Val Lys Leu Phe Glu Lys Trp Gly Trp Lys Lys Ile
            180                 185                 190

GCT ACC ATC CAG CAG ACC ACT GAG GTC TTC ACT TCG ACT CTG GAC GAC   624
Ala Thr Ile Gln Gln Thr Thr Glu Val Phe Thr Ser Thr Leu Asp Asp
        195                 200                 205

CTG GAG GAA CGA GTG AAG GAG GCT GGA ATT GAG ATT ACT TTC CGC CAG   672
Leu Glu Glu Arg Val Lys Glu Ala Gly Ile Glu Ile Thr Phe Arg Gln
    210                 215                 220

AGT TTC TTC TCA GAT CCA GCT GTG CCC GTC AAA AAC CTG AAG CGC CAG   720
Ser Phe Phe Ser Asp Pro Ala Val Pro Val Lys Asn Leu Lys Arg Gln
225                 230                 235                 240

GAT GCC CGA ATC ATC GTG GGA CTT TTC TAT GAG ACT GAA GCC CGG AAA   768
Asp Ala Arg Ile Ile Val Gly Leu Phe Tyr Glu Thr Glu Ala Arg Lys
                245                 250                 255

GTT TTT TGT GAG GTG TAC AAG GAG CGT CTC TTT GGG AAG AAG TAC GTC   816
Val Phe Cys Glu Val Tyr Lys Glu Arg Leu Phe Gly Lys Lys Tyr Val
            260                 265                 270

TGG TTC CTC ATT GGG TGG TAT GCT GAC AAT TGG TTC AAG ATC TAC GAC   864
Trp Phe Leu Ile Gly Trp Tyr Ala Asp Asn Trp Phe Lys Ile Tyr Asp
        275                 280                 285

CCT TCT ATC AAC TGC ACA GTG GAT GAG ATG ACT GAG GCG GTG GAG GGC   912
Pro Ser Ile Asn Cys Thr Val Asp Glu Met Thr Glu Ala Val Glu Gly
    290                 295                 300

CAC ATC ACA ACT GAG ATT GTC ATG CTG AAT CCT GCC AAT ACC CGC AGC   960
His Ile Thr Thr Glu Ile Val Met Leu Asn Pro Ala Asn Thr Arg Ser
305                 310                 315                 320

ATT TCC AAC ATG ACA TCC CAG GAA TTT GTG GAG AAA CTA ACC AAG CGA  1008
Ile Ser Asn Met Thr Ser Gln Glu Phe Val Glu Lys Leu Thr Lys Arg
                325                 330                 335
```

```
CTG AAA AGA CAC CCT GAG GAG ACA GGA GGC TTC CAG GAG GCA CCG CTG      1056
Leu Lys Arg His Pro Glu Glu Thr Gly Gly Phe Gln Glu Ala Pro Leu
            340                 345                 350

GCC TAT GAT GCC ATC TGG GCC TTG GCA CTG GCC CTG AAC AAG ACA TCT      1104
Ala Tyr Asp Ala Ile Trp Ala Leu Ala Leu Ala Leu Asn Lys Thr Ser
        355                 360                 365

GGA GGA GGC GGC CGT TCT GGT GTG CGC CTG GAG GAC TTC AAC TAC AAC      1152
Gly Gly Gly Gly Arg Ser Gly Val Arg Leu Glu Asp Phe Asn Tyr Asn
    370                 375                 380

AAC CAG ACC ATT ACC GAC CAA ATC TAC CGG GCA ATG AAC TCT TCG TCC      1200
Asn Gln Thr Ile Thr Asp Gln Ile Tyr Arg Ala Met Asn Ser Ser Ser
385                 390                 395                 400

TTT GAG GGT GTC TCT GGC CAT GTG GTG TTT GAT GCC AGC GGC TCT CGG      1248
Phe Glu Gly Val Ser Gly His Val Val Phe Asp Ala Ser Gly Ser Arg
                405                 410                 415

ATG GCA TGG ACG CTT ATC GAG CAG CTT CAG GGT GGC AGC TAC AAG AAG      1296
Met Ala Trp Thr Leu Ile Glu Gln Leu Gln Gly Gly Ser Tyr Lys Lys
            420                 425                 430

ATT GGC TAC TAT GAC AGC ACC AAG GAT GAT CTT TCC TGG TCC AAA ACA      1344
Ile Gly Tyr Tyr Asp Ser Thr Lys Asp Asp Leu Ser Trp Ser Lys Thr
        435                 440                 445

GAT AAA TGG ATT GGA GGG TCC CCC CCA GCT GAC CAG ACC CTG GTC ATC      1392
Asp Lys Trp Ile Gly Gly Ser Pro Pro Ala Asp Gln Thr Leu Val Ile
    450                 455                 460

AAG ACA TTC CGC TTC CTG TCA CAG AAA CTC TTT ATC TCC GTC TCA GTT      1440
Lys Thr Phe Arg Phe Leu Ser Gln Lys Leu Phe Ile Ser Val Ser Val
465                 470                 475                 480

CTC TCC AGC CTG GGC ATT GTC CTA GCT GTT GTC TGT CTG TCC TTT AAC      1488
Leu Ser Ser Leu Gly Ile Val Leu Ala Val Val Cys Leu Ser Phe Asn
                485                 490                 495

ATC TAC AAC TCA CAT GTC CGT TAT ATC CAG AAC TCA CAG CCC AAC CTG      1536
Ile Tyr Asn Ser His Val Arg Tyr Ile Gln Asn Ser Gln Pro Asn Leu
            500                 505                 510
```

```
AAC AAC CTG ACT GCT GTG GGC TGC TCA CTG GCT TTA GCT GCT GTC TTC      1584
Asn Asn Leu Thr Ala Val Gly Cys Ser Leu Ala Leu Ala Ala Val Phe
        515                 520                 525

CCC CTG GGG CTC GAT GGT TAC CAC ATT GGG AGG AAC CAG TTT CCT TTC      1632
Pro Leu Gly Leu Asp Gly Tyr His Ile Gly Arg Asn Gln Phe Pro Phe
    530                 535                 540

GTC TGC CAG GCC CGC CTC TGG CTC CTG GGC CTG GGC TTT AGT CTG GGC      1680
Val Cys Gln Ala Arg Leu Trp Leu Leu Gly Leu Gly Phe Ser Leu Gly
545                 550                 555                 560

TAC GGT TCC ATG TTC ACC AAG ATT TGG TGG GTC CAC ACG GTC TTC ACA      1728
Tyr Gly Ser Met Phe Thr Lys Ile Trp Trp Val His Thr Val Phe Thr
                565                 570                 575

AAG AAG GAA GAA AAG AAG GAG TGG AGG AAG ACT CTG GAA CCC TGG AAG      1776
Lys Lys Glu Glu Lys Lys Glu Trp Arg Lys Thr Leu Glu Pro Trp Lys
            580                 585                 590

CTG TAT GCC ACA GTG GGC CTG CTG GTG GGC ATG GAT GTC CTC ACT CTC      1824
Leu Tyr Ala Thr Val Gly Leu Leu Val Gly Met Asp Val Leu Thr Leu
        595                 600                 605

GCC ATC TGG CAG ATC GTG GAC CCT CTG CAC CGG ACC ATT GAG ACA TTT      1872
Ala Ile Trp Gln Ile Val Asp Pro Leu His Arg Thr Ile Glu Thr Phe
    610                 615                 620

GCC AAG GAG GAA CCT AAG GAA GAT ATT GAC GTC TCT ATT CTG CCC CAG      1920
Ala Lys Glu Glu Pro Lys Glu Asp Ile Asp Val Ser Ile Leu Pro Gln
625                 630                 635                 640

CTG GAG CAT TGC AGC TCC AGG AAG ATG AAT ACA TGG CTT GGC ATT TTC      1968
Leu Glu His Cys Ser Ser Arg Lys Met Asn Thr Trp Leu Gly Ile Phe
                645                 650                 655

TAT GGT TAC AAG GGG CTG CTG CTG CTG CTG GGA ATC TTC CTT GCT TAT      2016
Tyr Gly Tyr Lys Gly Leu Leu Leu Leu Leu Gly Ile Phe Leu Ala Tyr
            660                 665                 670

GAG ACC AAG AGT GTG TCC ACT GAG AAG ATC AAT GAT CAC CGG GCT GTG      2064
Glu Thr Lys Ser Val Ser Thr Glu Lys Ile Asn Asp His Arg Ala Val
```

675 680 685

GGC ATG GCT ATC TAC AAT GTG GCA GTC CTG TGC CTC ATC ACT GCT CCT 2112
Gly Met Ala Ile Tyr Asn Val Ala Val Leu Cys Leu Ile Thr Ala Pro
690 695 700

GTC ACC ATG ATT CTG TCC AGC CAG CAG GAT GCA GCC TTT GCC TTT GCC 2160
Val Thr Met Ile Leu Ser Ser Gln Gln Asp Ala Ala Phe Ala Phe Ala
705 710 715 720

TCT CTT GCC ATA GTT TTC TCC TCC TAT ATC ACT CTT GTT GTG CTC TTT 2208
Ser Leu Ala Ile Val Phe Ser Ser Tyr Ile Thr Leu Val Val Leu Phe
725 730 735

GTG CCC AAG ATG CGC AGG CTG ATC ACC CGA GGG GAA TGG CAG TCG GAG 2256
Val Pro Lys Met Arg Arg Leu Ile Thr Arg Gly Glu Trp Gln Ser Glu
740 745 750

GCG CAG GAC ACC ATG AAG ACA GGG TCA TCG ACC AAC AAC AAC GAG GAG 2304
Ala Gln Asp Thr Met Lys Thr Gly Ser Ser Thr Asn Asn Asn Glu Glu
755 760 765

GAG AAG TCC CGG CTG TTG GAG AAG GAG AAC CGT GAA CTG GAA AAG ATC 2352
Glu Lys Ser Arg Leu Leu Glu Lys Glu Asn Arg Glu Leu Glu Lys Ile
770 775 780

ATT GCT GAG AAA GAG GAG CGT GTC TCT GAA CTG CGC CAT CAA CTC CAG 2400
Ile Ala Glu Lys Glu Glu Arg Val Ser Glu Leu Arg His Gln Leu Gln
785 790 795 800

TCT CGG CAG CAG CTC CGC TCC CGG CGC CAC CCA CCG ACA CCC CCA GAA 2448
Ser Arg Gln Gln Leu Arg Ser Arg Arg His Pro Pro Thr Pro Pro Glu
805 810 815

CCC TCT GGG GGC CTG CCC AGG GGA CCC CCT GAG CCC CCC GAC CGG CTT- 2496
Pro Ser Gly Gly Leu Pro Arg Gly Pro Pro Glu Pro Pro Asp Arg Leu
820 825 830

AGC TGT GAT GGG AGT CGA GTG CAT TTG CTT TAT AAG TGA 2535
Ser Cys Asp Gly Ser Arg Val His Leu Leu Tyr Lys *
835 840

(2) INFORMATION FOR SEQ ID NO: 51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 844 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

```
Met Gly Pro Gly Ala Pro Phe Ala Arg Val Gly Trp Pro Leu Pro Leu
 1               5                  10                  15

Leu Val Val Met Ala Ala Gly Val Ala Pro Val Trp Ala Ser His Ser
            20                  25                  30

Pro His Leu Pro Arg Pro His Ser Arg Val Pro Pro His Pro Ser Ser
        35                  40                  45

Glu Arg Arg Ala Val Tyr Ile Gly Ala Leu Phe Pro Met Ser Gly Gly
    50                  55                  60

Trp Pro Gly Gly Gln Ala Cys Gln Pro Ala Val Glu Met Ala Leu Glu
65                  70                  75                  80

Asp Val Asn Ser Arg Arg Asp Ile Leu Pro Asp Tyr Glu Leu Lys Leu
                85                  90                  95

Ile His His Asp Ser Lys Cys Asp Pro Gly Gln Ala Thr Lys Tyr Leu
            100                 105                 110

Tyr Glu Leu Leu Tyr Asn Asp Pro Ile Lys Ile Ile Leu Met Pro Gly
        115                 120                 125

Cys Ser Ser Val Ser Thr Leu Val Ala Glu Ala Ala Arg Met Trp Asn
    130                 135                 140

Leu Ile Val Leu Ser Tyr Gly Ser Ser Ser Pro Ala Leu Ser Asn Arg
145                 150                 155                 160

Gln Arg Phe Pro Thr Phe Phe Arg Thr His Pro Ser Ala Thr Leu His
                165                 170                 175
```

```
Asn Pro Thr Arg Val Lys Leu Phe Glu Lys Trp Gly Trp Lys Lys Ile
            180                 185                 190

Ala Thr Ile Gln Gln Thr Thr Glu Val Phe Thr Ser Thr Leu Asp Asp
        195                 200                 205

Leu Glu Glu Arg Val Lys Glu Ala Gly Ile Glu Ile Thr Phe Arg Gln
    210                 215                 220

Ser Phe Phe Ser Asp Pro Ala Val Pro Val Lys Asn Leu Lys Arg Gln
225                 230                 235                 240

Asp Ala Arg Ile Ile Val Gly Leu Phe Tyr Glu Thr Glu Ala Arg Lys
                245                 250                 255

Val Phe Cys Glu Val Tyr Lys Glu Arg Leu Phe Gly Lys Lys Tyr Val
            260                 265                 270

Trp Phe Leu Ile Gly Trp Tyr Ala Asp Asn Trp Phe Lys Ile Tyr Asp
        275                 280                 285

Pro Ser Ile Asn Cys Thr Val Asp Glu Met Thr Glu Ala Val Glu Gly
    290                 295                 300

His Ile Thr Thr Glu Ile Val Met Leu Asn Pro Ala Asn Thr Arg Ser
305                 310                 315                 320

Ile Ser Asn Met Thr Ser Gln Glu Phe Val Glu Lys Leu Thr Lys Arg
                325                 330                 335

Leu Lys Arg His Pro Glu Glu Thr Gly Gly Phe Gln Glu Ala Pro Leu
            340                 345                 350

Ala Tyr Asp Ala Ile Trp Ala Leu Ala Leu Ala Leu Asn Lys Thr Ser
        355                 360                 365

Gly Gly Gly Gly Arg Ser Gly Val Arg Leu Glu Asp Phe Asn Tyr Asn
    370                 375                 380

Asn Gln Thr Ile Thr Asp Gln Ile Tyr Arg Ala Met Asn Ser Ser Ser
385                 390                 395                 400
```

Phe Glu Gly Val Ser Gly His Val Val Phe Asp Ala Ser Gly Ser Arg
405 410 415

Met Ala Trp Thr Leu Ile Glu Gln Leu Gln Gly Gly Ser Tyr Lys Lys
420 425 430

Ile Gly Tyr Tyr Asp Ser Thr Lys Asp Asp Leu Ser Trp Ser Lys Thr
435 440 445

Asp Lys Trp Ile Gly Gly Ser Pro Pro Ala Asp Gln Thr Leu Val Ile
450 455 460

Lys Thr Phe Arg Phe Leu Ser Gln Lys Leu Phe Ile Ser Val Ser Val
465 470 475 480

Leu Ser Ser Leu Gly Ile Val Leu Ala Val Val Cys Leu Ser Phe Asn
485 490 495

Ile Tyr Asn Ser His Val Arg Tyr Ile Gln Asn Ser Gln Pro Asn Leu
500 505 510

Asn Asn Leu Thr Ala Val Gly Cys Ser Leu Ala Leu Ala Ala Val Phe
515 520 525

Pro Leu Gly Leu Asp Gly Tyr His Ile Gly Arg Asn Gln Phe Pro Phe
530 535 540

Val Cys Gln Ala Arg Leu Trp Leu Leu Gly Leu Gly Phe Ser Leu Gly
545 550 555 560

Tyr Gly Ser Met Phe Thr Lys Ile Trp Trp Val His Thr Val Phe Thr
565 570 575

Lys Lys Glu Glu Lys Lys Glu Trp Arg Lys Thr Leu Glu Pro Trp Lys
580 585 590

Leu Tyr Ala Thr Val Gly Leu Leu Val Gly Met Asp Val Leu Thr Leu
595 600 605

Ala Ile Trp Gln Ile Val Asp Pro Leu His Arg Thr Ile Glu Thr Phe
610 615 620

Ala Lys Glu Glu Pro Lys Glu Asp Ile Asp Val Ser Ile Leu Pro Gln

625 630 635 640

Leu Glu His Cys Ser Ser Arg Lys Met Asn Thr Trp Leu Gly Ile Phe
645 650 655

Tyr Gly Tyr Lys Gly Leu Leu Leu Leu Leu Gly Ile Phe Leu Ala Tyr
660 665 670

Glu Thr Lys Ser Val Ser Thr Glu Lys Ile Asn Asp His Arg Ala Val
675 680 685

Gly Met Ala Ile Tyr Asn Val Ala Val Leu Cys Leu Ile Thr Ala Pro
690 695 700

Val Thr Met Ile Leu Ser Ser Gln Gln Asp Ala Ala Phe Ala Phe Ala
705 710 715 720

Ser Leu Ala Ile Val Phe Ser Ser Tyr Ile Thr Leu Val Val Leu Phe
725 730 735

Val Pro Lys Met Arg Arg Leu Ile Thr Arg Gly Glu Trp Gln Ser Glu
740 745 750

Ala Gln Asp Thr Met Lys Thr Gly Ser Ser Thr Asn Asn Asn Glu Glu
755 760 765

Glu Lys Ser Arg Leu Leu Glu Lys Glu Asn Arg Glu Leu Glu Lys Ile
770 775 780

Ile Ala Glu Lys Glu Glu Arg Val Ser Glu Leu Arg His Gln Leu Gln
785 790 795 800

Ser Arg Gln Gln Leu Arg Ser Arg Arg His Pro Pro Thr Pro Pro Glu
805 810 815

Pro Ser Gly Gly Leu Pro Arg Gly Pro Pro Glu Pro Pro Asp Arg Leu
820 825 830

Ser Cys Asp Gly Ser Arg Val His Leu Leu Tyr Lys
835 840

(2) INFORMATION FOR SEQ ID NO: 52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2895 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Canis familiaris

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..2895

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

```
ATG CTG CTG CTG CTC CTG CCG CTG GCG CTG GCG CCG CTC TTC CTC CGC      48
Met Leu Leu Leu Leu Leu Pro Leu Ala Leu Ala Pro Leu Phe Leu Arg
 1               5                  10                  15

CCC CCG GGC GCG GGC GGG GCA CAG ACC CCC AAC GCC ACC TCG GAA GGT      96
Pro Pro Gly Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly
            20                  25                  30

TGC CAG ATC ATA CAC CCG CCT TGG GAA GGG GGT ATC AGG TAC AGG GGC     144
Cys Gln Ile Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly
        35                  40                  45

CTG ACT CGT GAC CAG GTG AAG GCT ATC AAC TTC CTG CCG GTG GAC TAT     192
Leu Thr Arg Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr
    50                  55                  60

GAG ATT GAG TAT GTG TGC CGG GGA GAG CGA GAG GTG GTG GGG CCC AAG     240
Glu Ile Glu Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys
65                  70                  75                  80

GTC CGA AAG TGC CTG GCC AAT GGC TCC TGG ACA GAT ATG GAC ACA CCC     288
```

```
Val Arg Lys Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro
                85                  90                  95

AGC CGC TGT GTC CGA ATC TGT TCC AAG TCA TAT TTG GCC CTG GAA AAT   336
Ser Arg Cys Val Arg Ile Cys Ser Lys Ser Tyr Leu Ala Leu Glu Asn
            100                 105                 110

GGG AAG GTC TTC CTG ACG GGT GGG GAC CTC CCC GCT CTG GAT GGA GCC   384
Gly Lys Val Phe Leu Thr Gly Gly Asp Leu Pro Ala Leu Asp Gly Ala
        115                 120                 125

CGG GTG GAT TTC CGG TGT GAC CCT GAC TTC CAT CTT GTG GGC AGC TCC   432
Arg Val Asp Phe Arg Cys Asp Pro Asp Phe His Leu Val Gly Ser Ser
    130                 135                 140

CGG AGT ATC TGT AGT CAG GGC CAG TGG AGC ACT CCC AAG CCC CAC TGC   480
Arg Ser Ile Cys Ser Gln Gly Gln Trp Ser Thr Pro Lys Pro His Cys
145                 150                 155                 160

CAG GTG AGC CGA ACG CCG CAC TCA GAG CGG CGA GCG GTG TAC ATC GGG   528
Gln Val Ser Arg Thr Pro His Ser Glu Arg Arg Ala Val Tyr Ile Gly
                165                 170                 175

GCG CTG TTT CCC ATG AGC GGG GGC TGG CCG GGG GGC CAG GCC TGC CAG   576
Ala Leu Phe Pro Met Ser Gly Gly Trp Pro Gly Gly Gln Ala Cys Gln
            180                 185                 190

CCC GCG GTG GAG ATG GCG CTG GAG GAC GTG AAT AGC CGC AGG GAC ATC   624
Pro Ala Val Glu Met Ala Leu Glu Asp Val Asn Ser Arg Arg Asp Ile
        195                 200                 205

CTG CCG GAC TAC GAG CTC AAG CTC ATC CAC CAC GAC AGC AAG TGT GAC   672
Leu Pro Asp Tyr Glu Leu Lys Leu Ile His His Asp Ser Lys Cys Asp
    210                 215                 220

CCA GGC CAA GCT ACC AAG TAC CTG TAT GAA CTG CTC TAC AAC GAC CCC   720
Pro Gly Gln Ala Thr Lys Tyr Leu Tyr Glu Leu Leu Tyr Asn Asp Pro
225                 230                 235                 240

ATC AAG ATC ATC CTC ATG CCT GGC TGC AGC TCT GTC TCC ACG CTT GTG   768
Ile Lys Ile Ile Leu Met Pro Gly Cys Ser Ser Val Ser Thr Leu Val
                245                 250                 255
```

```
GCT GAG GCT GCC AGG ATG TGG AAC CTC ATT GTG CTC TCC TAT GGT TCC      816
Ala Glu Ala Ala Arg Met Trp Asn Leu Ile Val Leu Ser Tyr Gly Ser
            260                 265                 270

AGC TCA CCA GCT CTG TCC AAC CGG CAG CGC TTT CCT ACC TTC TTC CGA      864
Ser Ser Pro Ala Leu Ser Asn Arg Gln Arg Phe Pro Thr Phe Phe Arg
        275                 280                 285

ACT CAT CCC TCG GCC ACG CTC CAC AAC CCT ACG CGA GTG AAG CTC TTT      912
Thr His Pro Ser Ala Thr Leu His Asn Pro Thr Arg Val Lys Leu Phe
    290                 295                 300

GAG AAG TGG GGC TGG AGG AAG ATT GCC ACC ATC CAG CAG ACC ACC GAG      960
Glu Lys Trp Gly Trp Arg Lys Ile Ala Thr Ile Gln Gln Thr Thr Glu
305                 310                 315                 320

GTG TTC ACA TCG ACT CTG GAC GAC CTA GAG GAA CGA GTG AAG GAG GCT     1008
Val Phe Thr Ser Thr Leu Asp Asp Leu Glu Glu Arg Val Lys Glu Ala
                325                 330                 335

GGG ATT GAG ATT ACT TTC CGC CAG AGC TTC TTC TCA GAT CCT GCC GTG     1056
Gly Ile Glu Ile Thr Phe Arg Gln Ser Phe Phe Ser Asp Pro Ala Val
            340                 345                 350

CCT GTC AAG AAC CTC AAG CGC CAG GAT GCC CGA ATC ATC GTG GGA CTT     1104
Pro Val Lys Asn Leu Lys Arg Gln Asp Ala Arg Ile Ile Val Gly Leu
        355                 360                 365

TTC TAT GAG ACT GAA GCC CGG AAA GTG TTC TGT GAG GTA TAC AAG GAG     1152
Phe Tyr Glu Thr Glu Ala Arg Lys Val Phe Cys Glu Val Tyr Lys Glu
    370                 375                 380

CGG CTC TTT GGG AAG AAG TAT GTG TGG TTC CTC ATT GGG TGG TAT GCT     1200
Arg Leu Phe Gly Lys Lys Tyr Val Trp Phe Leu Ile Gly Trp Tyr Ala
385                 390                 395                 400

GAC AAT TGG TTC AAG ACC TAC GAC CCC TCC ATC AAC TGC ACA GTG GAT     1248
Asp Asn Trp Phe Lys Thr Tyr Asp Pro Ser Ile Asn Cys Thr Val Asp
                405                 410                 415

GAG ATG ACC GAG GCT GTG GAA GGC CAC ATC ACC ACT GAG ATT GTC ATG     1296
Glu Met Thr Glu Ala Val Glu Gly His Ile Thr Thr Glu Ile Val Met
            420                 425                 430
```

```
CTG AAC CCA GCC AAC ACC CGC AGC ATC TCC AAC ATG ACA TCC CAG GAG      1344
Leu Asn Pro Ala Asn Thr Arg Ser Ile Ser Asn Met Thr Ser Gln Glu
        435                 440                 445

TTT GTG GAG AAA CTG ACC AAG AGA CTC AAG AGA CAC CCT GAG GAG ACA      1392
Phe Val Glu Lys Leu Thr Lys Arg Leu Lys Arg His Pro Glu Glu Thr
    450                 455                 460

GGC GGC TTC CAG GAG GCA CCG CTG GCC TAT GAT GCC ATC TGG GCC TTG      1440
Gly Gly Phe Gln Glu Ala Pro Leu Ala Tyr Asp Ala Ile Trp Ala Leu
465                 470                 475                 480

GCA TTG GCC CTG AAC AAG ACA TCT GGA GGG AGC GGC CGT TCG GGG GTG      1488
Ala Leu Ala Leu Asn Lys Thr Ser Gly Gly Ser Gly Arg Ser Gly Val
                485                 490                 495

CGC CTG GAA GAC TTC AAC TAC AAC AAC CAG ACG ATC ACA GAC CAA ATC      1536
Arg Leu Glu Asp Phe Asn Tyr Asn Asn Gln Thr Ile Thr Asp Gln Ile
            500                 505                 510

TAC CGC GCA ATG AAC TCC TCG TCC TTT GAG GGT GTC TCT GGC CAC GTG      1584
Tyr Arg Ala Met Asn Ser Ser Ser Phe Glu Gly Val Ser Gly His Val
        515                 520                 525

GTG TTT GAT GCC AGC GGC TCA CGG ATG GCC TGG ACT CTG ATT GAG CAG      1632
Val Phe Asp Ala Ser Gly Ser Arg Met Ala Trp Thr Leu Ile Glu Gln
    530                 535                 540

CTG CAG GGT GGC AGC TAC AAG AAG ATC GGC TAC TAT GAC AGC ACC AAG      1680
Leu Gln Gly Gly Ser Tyr Lys Lys Ile Gly Tyr Tyr Asp Ser Thr Lys
545                 550                 555                 560

GAT GAC CTT TCC TGG TCT AAA ACG GAC AAA TGG ATT GGA GGG GCC CCC      1728
Asp Asp Leu Ser Trp Ser Lys Thr Asp Lys Trp Ile Gly Gly Ala Pro
                565                 570                 575

CCG GCC GAC CAG ACC CTG GTC ATC AAG ACA TTT CGC TTC ATG TCA CAG      1776
Pro Ala Asp Gln Thr Leu Val Ile Lys Thr Phe Arg Phe Met Ser Gln
            580                 585                 590

AAG CTC TTC ATT TCA GTC TCT GTC CTC TCC AGC CTG GGC ATT GTC CTG      1824
Lys Leu Phe Ile Ser Val Ser Val Leu Ser Ser Leu Gly Ile Val Leu
```

595 600 605

GCT GTG GTC TGT CTG TCC TTT AAC ATC TAC AAC TCT CAT GTC CGT TAC 1872
Ala Val Val Cys Leu Ser Phe Asn Ile Tyr Asn Ser His Val Arg Tyr
610 615 620

ATC CAG AAC TCC CAG CCC AAC TTG AAC AAT CTG ACT GCT GTG GGC TGC 1920
Ile Gln Asn Ser Gln Pro Asn Leu Asn Asn Leu Thr Ala Val Gly Cys
625 630 635 640

TCC CTG GCA TTG GCT GCC GTC TTC CCC CTG GGG CTA GAT GGG TAC CAC 1968
Ser Leu Ala Leu Ala Ala Val Phe Pro Leu Gly Leu Asp Gly Tyr His
645 650 655

ATC GGG AGA AGC CAG TTT CCT TTT GTG TGT CAG GCA CGC CTC TGG CTC 2016
Ile Gly Arg Ser Gln Phe Pro Phe Val Cys Gln Ala Arg Leu Trp Leu
660 665 670

CTG GGT CTG GGC TTC AGT CTG GGC TAT GGC TCC ATG TTC ACG AAG ATC 2064
Leu Gly Leu Gly Phe Ser Leu Gly Tyr Gly Ser Met Phe Thr Lys Ile
675 680 685

TGG TGG GTC CAC ACG GTC TTC ACT AAG AAG GAG GAG AAG AAG GAG TGG 2112
Trp Trp Val His Thr Val Phe Thr Lys Lys Glu Glu Lys Lys Glu Trp
690 695 700

AGG AAG ACC CTG GAG CCC TGG AAG CTG TAC ACC ACA GTG GGC TTG CTA 2160
Arg Lys Thr Leu Glu Pro Trp Lys Leu Tyr Thr Thr Val Gly Leu Leu
705 710 715 720

GTG GGC ATG GAT GTC CTC ACT CTT GCC ATT TGG CAG ATG GTA GAC CCC 2208
Val Gly Met Asp Val Leu Thr Leu Ala Ile Trp Gln Met Val Asp Pro
725 730 735

TTG CAC CGG ACC ATT GAG ACT TTT GCC AAG GAG GAA CCA AAG GAA GAT 2256
Leu His Arg Thr Ile Glu Thr Phe Ala Lys Glu Glu Pro Lys Glu Asp
740 745 750

ATT GAT GTG TCC ATC CTG CCC CAG CTG GAG CAC TGC AGC TCC AAG AAA 2304
Ile Asp Val Ser Ile Leu Pro Gln Leu Glu His Cys Ser Ser Lys Lys
755 760 765

ATG AAC ACC TGG CTT GGC ATT TTC TAT GGT TAC AAG GGG CTG CTG CTG 2352

```
Met Asn Thr Trp Leu Gly Ile Phe Tyr Gly Tyr Lys Gly Leu Leu Leu
 770                 775                 780

CTG CTA GGC ATC TTT CTT GCT TAT GAG ACC AAG AGC GTG TCT ACT GAG      2400
Leu Leu Gly Ile Phe Leu Ala Tyr Glu Thr Lys Ser Val Ser Thr Glu
785                 790                 795                 800

AAG ATC AAT GAC CAC CGG GCT GTG GGC ATG GCC ATG TAC AAC GTG GCG      2448
Lys Ile Asn Asp His Arg Ala Val Gly Met Ala Met Tyr Asn Val Ala
                805                 810                 815

GTC CTG TGC CTC ATC ACT GCC CCG GTC ACC ATG ATC CTG TCC AGC CAG      2496
Val Leu Cys Leu Ile Thr Ala Pro Val Thr Met Ile Leu Ser Ser Gln
            820                 825                 830

CAG GAT GCA GCT TTC GCC TTT GCA GCT CTT GCC ATA GTG TTC TCC TCC      2544
Gln Asp Ala Ala Phe Ala Phe Ala Ala Leu Ala Ile Val Phe Ser Ser
        835                 840                 845

TAC ATC ACT CTG GTC GTT CTG TTC GTG CCG AAG ATG CGC AGG TTG ATC      2592
Tyr Ile Thr Leu Val Val Leu Phe Val Pro Lys Met Arg Arg Leu Ile
 850                 855                 860

ACC CGG GGT GAG TGG CAG TCG GAG GCG CAG GAT ACC ATG AAA ACG GGG      2640
Thr Arg Gly Glu Trp Gln Ser Glu Ala Gln Asp Thr Met Lys Thr Gly
865                 870                 875                 880

TCG TCG ACC AAC AAC AAT GAG GAA GAG AAG TCC CGA CTG TTG GAG AAG      2688
Ser Ser Thr Asn Asn Asn Glu Glu Glu Lys Ser Arg Leu Leu Glu Lys
                885                 890                 895

GAG AAC CGG GAG CTG GAG AAG ATC ATT GCT GAG AAA GAG GAG CGA GTG      2736
Glu Asn Arg Glu Leu Glu Lys Ile Ile Ala Glu Lys Glu Glu Arg Val
            900                 905                 910

TCC GAG CTG CGC CAT CAG CTT CGT TCT CGG CAG CAG CTG CGC CCT CGG      2784
Ser Glu Leu Arg His Gln Leu Arg Ser Arg Gln Gln Leu Arg Pro Arg
        915                 920                 925

CGT CAC CCC CCG ACG CCC CCA GAC CCC TCA GGG GGC CTG CCC AGG GGA      2832
Arg His Pro Pro Thr Pro Pro Asp Pro Ser Gly Gly Leu Pro Arg Gly
 930                 935                 940
```

CCC CAT GAG CCC CCT GAC CGG CTC AGC TGT GAC GGG AGC CGG GTT CAC 2880
Pro His Glu Pro Pro Asp Arg Leu Ser Cys Asp Gly Ser Arg Val His
945 950 955 960

TTG CTG TAC AAG TGA 2895
Leu Leu Tyr Lys *
965

(2) INFORMATION FOR SEQ ID NO: 53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 964 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

Met Leu Leu Leu Leu Leu Pro Leu Ala Leu Ala Pro Leu Phe Leu Arg
1 5 10 15

Pro Pro Gly Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly
20 25 30

Cys Gln Ile Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly
35 40 45

Leu Thr Arg Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr
50 55 60

Glu Ile Glu Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys
65 70 75 80

Val Arg Lys Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro
85 90 95

Ser Arg Cys Val Arg Ile Cys Ser Lys Ser Tyr Leu Ala Leu Glu Asn
100 105 110

Gly Lys Val Phe Leu Thr Gly Gly Asp Leu Pro Ala Leu Asp Gly Ala
115 120 125

```
Arg Val Asp Phe Arg Cys Asp Pro Asp Phe His Leu Val Gly Ser Ser
    130                 135                 140

Arg Ser Ile Cys Ser Gln Gly Gln Trp Ser Thr Pro Lys Pro His Cys
145                 150                 155                 160

Gln Val Ser Arg Thr Pro His Ser Glu Arg Arg Ala Val Tyr Ile Gly
                165                 170                 175

Ala Leu Phe Pro Met Ser Gly Gly Trp Pro Gly Gly Gln Ala Cys Gln
            180                 185                 190

Pro Ala Val Glu Met Ala Leu Glu Asp Val Asn Ser Arg Arg Asp Ile
        195                 200                 205

Leu Pro Asp Tyr Glu Leu Lys Leu Ile His His Asp Ser Lys Cys Asp
    210                 215                 220

Pro Gly Gln Ala Thr Lys Tyr Leu Tyr Glu Leu Leu Tyr Asn Asp Pro
225                 230                 235                 240

Ile Lys Ile Ile Leu Met Pro Gly Cys Ser Ser Val Ser Thr Leu Val
                245                 250                 255

Ala Glu Ala Ala Arg Met Trp Asn Leu Ile Val Leu Ser Tyr Gly Ser
            260                 265                 270

Ser Ser Pro Ala Leu Ser Asn Arg Gln Arg Phe Pro Thr Phe Phe Arg
        275                 280                 285

Thr His Pro Ser Ala Thr Leu His Asn Pro Thr Arg Val Lys Leu Phe
    290                 295                 300

Glu Lys Trp Gly Trp Arg Lys Ile Ala Thr Ile Gln Gln Thr Thr Glu
305                 310                 315                 320

Val Phe Thr Ser Thr Leu Asp Asp Leu Glu Glu Arg Val Lys Glu Ala
                325                 330                 335

Gly Ile Glu Ile Thr Phe Arg Gln Ser Phe Phe Ser Asp Pro Ala Val
            340                 345                 350
```

Pro Val Lys Asn Leu Lys Arg Gln Asp Ala Arg Ile Ile Val Gly Leu
355 360 365

Phe Tyr Glu Thr Glu Ala Arg Lys Val Phe Cys Glu Val Tyr Lys Glu
370 375 380

Arg Leu Phe Gly Lys Lys Tyr Val Trp Phe Leu Ile Gly Trp Tyr Ala
385 390 395 400

Asp Asn Trp Phe Lys Thr Tyr Asp Pro Ser Ile Asn Cys Thr Val Asp
405 410 415

Glu Met Thr Glu Ala Val Glu Gly His Ile Thr Thr Glu Ile Val Met
420 425 430

Leu Asn Pro Ala Asn Thr Arg Ser Ile Ser Asn Met Thr Ser Gln Glu
435 440 445

Phe Val Glu Lys Leu Thr Lys Arg Leu Lys Arg His Pro Glu Glu Thr
450 455 460

Gly Gly Phe Gln Glu Ala Pro Leu Ala Tyr Asp Ala Ile Trp Ala Leu
465 470 475 480

Ala Leu Ala Leu Asn Lys Thr Ser Gly Gly Ser Gly Arg Ser Gly Val
485 490 495

Arg Leu Glu Asp Phe Asn Tyr Asn Asn Gln Thr Ile Thr Asp Gln Ile
500 505 510

Tyr Arg Ala Met Asn Ser Ser Ser Phe Glu Gly Val Ser Gly His Val
515 520 525

Val Phe Asp Ala Ser Gly Ser Arg Met Ala Trp Thr Leu Ile Glu Gln
530 535 540

Leu Gln Gly Gly Ser Tyr Lys Lys Ile Gly Tyr Tyr Asp Ser Thr Lys
545 550 555 560

Asp Asp Leu Ser Trp Ser Lys Thr Asp Lys Trp Ile Gly Gly Ala Pro
565 570 575

Pro Ala Asp Gln Thr Leu Val Ile Lys Thr Phe Arg Phe Met Ser Gln

580 585 590

Lys Leu Phe Ile Ser Val Ser Val Leu Ser Ser Leu Gly Ile Val Leu
595 600 605

Ala Val Val Cys Leu Ser Phe Asn Ile Tyr Asn Ser His Val Arg Tyr
610 615 620

Ile Gln Asn Ser Gln Pro Asn Leu Asn Asn Leu Thr Ala Val Gly Cys
625 630 635 640

Ser Leu Ala Leu Ala Ala Val Phe Pro Leu Gly Leu Asp Gly Tyr His
645 650 655

Ile Gly Arg Ser Gln Phe Pro Phe Val Cys Gln Ala Arg Leu Trp Leu
660 665 670

Leu Gly Leu Gly Phe Ser Leu Gly Tyr Gly Ser Met Phe Thr Lys Ile
675 680 685

Trp Trp Val His Thr Val Phe Thr Lys Lys Glu Glu Lys Lys Glu Trp
690 695 700

Arg Lys Thr Leu Glu Pro Trp Lys Leu Tyr Thr Thr Val Gly Leu Leu
705 710 715 720

Val Gly Met Asp Val Leu Thr Leu Ala Ile Trp Gln Met Val Asp Pro
725 730 735

Leu His Arg Thr Ile Glu Thr Phe Ala Lys Glu Glu Pro Lys Glu Asp
740 745 750

Ile Asp Val Ser Ile Leu Pro Gln Leu Glu His Cys Ser Ser Lys Lys
755 760 765

Met Asn Thr Trp Leu Gly Ile Phe Tyr Gly Tyr Lys Gly Leu Leu Leu
770 775 780

Leu Leu Gly Ile Phe Leu Ala Tyr Glu Thr Lys Ser Val Ser Thr Glu
785 790 795 800

Lys Ile Asn Asp His Arg Ala Val Gly Met Ala Met Tyr Asn Val Ala
805 810 815

Val Leu Cys Leu Ile Thr Ala Pro Val Thr Met Ile Leu Ser Ser Gln
820 825 830

Gln Asp Ala Ala Phe Ala Phe Ala Ala Leu Ala Ile Val Phe Ser Ser
835 840 845

Tyr Ile Thr Leu Val Val Leu Phe Val Pro Lys Met Arg Arg Leu Ile
850 855 860

Thr Arg Gly Glu Trp Gln Ser Glu Ala Gln Asp Thr Met Lys Thr Gly
865 870 875 880

Ser Ser Thr Asn Asn Asn Glu Glu Glu Lys Ser Arg Leu Leu Glu Lys
885 890 895

Glu Asn Arg Glu Leu Glu Lys Ile Ile Ala Glu Lys Glu Glu Arg Val
900 905 910

Ser Glu Leu Arg His Gln Leu Arg Ser Arg Gln Gln Leu Arg Pro Arg
915 920 925

Arg His Pro Pro Thr Pro Pro Asp Pro Ser Gly Gly Leu Pro Arg Gly
930 935 940

Pro His Glu Pro Pro Asp Arg Leu Ser Cys Asp Gly Ser Arg Val His
945 950 955 960

Leu Leu Tyr Lys

(2) INFORMATION FOR SEQ ID NO: 54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1737 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..1737

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

```
ATG TTG CTG CTG CTG CTA CTG GCG CCA CTC TTC CTC CGC CCC CCG GGC      48
Met Leu Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
  1               5                  10                  15

GCG GGC GGG GCG CAG ACC CCC AAC GCC ACC TCA GAA GGT TGC CAG ATC      96
Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
            20                  25                  30

ATA CAC CCG CCC TGG GAA GGG GGC ATC AGG TAC CGG GGC CTG ACT CGG     144
Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
        35                  40                  45

GAC CAG GTG AAG GCT ATC AAC TTC CTG CCA GTG GAC TAT GAG ATT GAG     192
Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu
    50                  55                  60

TAT GTG TGC CGG GGG GAG CGC GAG GTG GTG GGG CCC AAG GTC CGC AAG     240
Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys
65                  70                  75                  80

TGC CTG GCC AAC GGC TCC TGG ACA GAT ATG GAC ACA CCC AGC CGC TGT     288
Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys
                85                  90                  95

GTC CGA ATC TGC TCC AAG TCT TAT TTG ACC CTG GAA AAT GGG AAG GTT     336
Val Arg Ile Cys Ser Lys Ser Tyr Leu Thr Leu Glu Asn Gly Lys Val
            100                 105                 110

TTC CTG ACG GGT GGG GAC CTC CCA GCT CTG GAC GGA GCC CGG GTG GAT     384
Phe Leu Thr Gly Gly Asp Leu Pro Ala Leu Asp Gly Ala Arg Val Asp
        115                 120                 125
```

```
TTC CGG TGT GAC CCC GAC TTC CAT CTG GTG GGC AGC TCC CGG AGC ATC        432
Phe Arg Cys Asp Pro Asp Phe His Leu Val Gly Ser Ser Arg Ser Ile
    130                 135                 140

TGT AGT CAG GGC CAG TGG AGC ACC CCC AAG CCC CAC TGC CAG GTG AAT        480
Cys Ser Gln Gly Gln Trp Ser Thr Pro Lys Pro His Cys Gln Val Asn
145                 150                 155                 160

CGA ACG CCA CAC TCA GAA CGG CGC GCA GTG TAC ATC GGG GCA CTG TTT        528
Arg Thr Pro His Ser Glu Arg Arg Ala Val Tyr Ile Gly Ala Leu Phe
                165                 170                 175

CCC ATG AGC GGG GGC TGG CCA GGG GGC CAG GCC TGC CAG CCC GCG GTG        576
Pro Met Ser Gly Gly Trp Pro Gly Gly Gln Ala Cys Gln Pro Ala Val
            180                 185                 190

GAG ATG GCG CTG GAG GAC GTG AAT AGC CGC AGG GAC ATC CTG CCG GAC        624
Glu Met Ala Leu Glu Asp Val Asn Ser Arg Arg Asp Ile Leu Pro Asp
        195                 200                 205

TAT GAG CTC AAG CTC ATC CAC CAC GAC AGC AAG TGT GAT CCA GGC CAA        672
Tyr Glu Leu Lys Leu Ile His His Asp Ser Lys Cys Asp Pro Gly Gln
    210                 215                 220

GCC ACC AAG TAC CTA TAT GAG CTG CTC TAC AAC GAC CCT ATC AAG ATC        720
Ala Thr Lys Tyr Leu Tyr Glu Leu Leu Tyr Asn Asp Pro Ile Lys Ile
225                 230                 235                 240

ATC CTT ATG CCT GGC TGC AGC TCT GTC TCC ACG CTG GTG GCT GAG GCT        768
Ile Leu Met Pro Gly Cys Ser Ser Val Ser Thr Leu Val Ala Glu Ala
                245                 250                 255

GCT AGG ATG TGG AAC CTC ATT GTG CTT TCC TAT GGC TCC AGC TCA CCA        816
Ala Arg Met Trp Asn Leu Ile Val Leu Ser Tyr Gly Ser Ser Ser Pro
            260                 265                 270

GCC CTG TCA AAC CGG CAG CGT TTC CCC ACT TTC TTC CGA ACG CAC CCA        864
Ala Leu Ser Asn Arg Gln Arg Phe Pro Thr Phe Phe Arg Thr His Pro
        275                 280                 285

TCA GCC ACA CTC CAC AAC CCT ACC CGC GTG AAA CTC TTT GAA AAG TGG        912
Ser Ala Thr Leu His Asn Pro Thr Arg Val Lys Leu Phe Glu Lys Trp
    290                 295                 300
```

```
GGC TGG AAG AAG ATT GCT ACC ATC CAG CAG ACC ACT GAG GTC TTC ACT      960
Gly Trp Lys Lys Ile Ala Thr Ile Gln Gln Thr Thr Glu Val Phe Thr
305                 310                 315                 320

TCG ACT CTG GAC GAC CTG GAG GAA CGA GTG AAG GAG GCT GGA ATT GAG     1008
Ser Thr Leu Asp Asp Leu Glu Glu Arg Val Lys Glu Ala Gly Ile Glu
                325                 330                 335

ATT ACT TTC CGC CAG AGT TTC TTC TCA GAT CCA GCT GTG CCC GTC AAA     1056
Ile Thr Phe Arg Gln Ser Phe Phe Ser Asp Pro Ala Val Pro Val Lys
            340                 345                 350

AAC CTG AAG CGC CAG GAT GCC CGA ATC ATC GTG GGA CTT TTC TAT GAG     1104
Asn Leu Lys Arg Gln Asp Ala Arg Ile Ile Val Gly Leu Phe Tyr Glu
        355                 360                 365

ACT GAA GCC CGG AAA GTT TTT TGT GAG GTG TAC AAG GAG CGT CTC TTT     1152
Thr Glu Ala Arg Lys Val Phe Cys Glu Val Tyr Lys Glu Arg Leu Phe
    370                 375                 380

GGG AAG AAG TAC GTC TGG TTC CTC ATT GGG TGG TAT GCT GAC AAT TGG     1200
Gly Lys Lys Tyr Val Trp Phe Leu Ile Gly Trp Tyr Ala Asp Asn Trp
385                 390                 395                 400

TTC AAG ATC TAC GAC CCT TCT ATC AAC TGC ACA GTG GAT GAG ATG ACT     1248
Phe Lys Ile Tyr Asp Pro Ser Ile Asn Cys Thr Val Asp Glu Met Thr
                405                 410                 415

GAG GCG GTG GAG GGC CAC ATC ACA ACT GAG ATT GTC ATG CTG AAT CCT     1296
Glu Ala Val Glu Gly His Ile Thr Thr Glu Ile Val Met Leu Asn Pro
            420                 425                 430

GCC AAT ACC CGC AGC ATT TCC AAC ATG ACA TCC CAG GAA TTT GTG GAG     1344
Ala Asn Thr Arg Ser Ile Ser Asn Met Thr Ser Gln Glu Phe Val Glu
        435                 440                 445

AAA CTA ACC AAG CGA CTG AAA AGA CAC CCT GAG GAG ACA GGA GGC TTC     1392
Lys Leu Thr Lys Arg Leu Lys Arg His Pro Glu Glu Thr Gly Gly Phe
    450                 455                 460

CAG GAG GCA CCG CTG GCC TAT GAT GCC ATC TGG GCC TTG GCA CTG GCC     1440
Gln Glu Ala Pro Leu Ala Tyr Asp Ala Ile Trp Ala Leu Ala Leu Ala
```

```
465                 470                 475                 480

CTG AAC AAG ACA TCT GGA GGA GGC GGC CGT TCT GGT GTG CGC CTG GAG      1488
Leu Asn Lys Thr Ser Gly Gly Gly Gly Arg Ser Gly Val Arg Leu Glu
                485                 490                 495

GAC TTC AAC TAC AAC AAC CAG ACC ATT ACC GAC CAA ATC TAC CGG GCA      1536
Asp Phe Asn Tyr Asn Asn Gln Thr Ile Thr Asp Gln Ile Tyr Arg Ala
            500                 505                 510

ATG AAC TCT TCG TCC TTT GAG GGT GTC TCT GGC CAT GTG GTG TTT GAT      1584
Met Asn Ser Ser Ser Phe Glu Gly Val Ser Gly His Val Val Phe Asp
        515                 520                 525

GCC AGC GGC TCT CGG ATG GCA TGG ACG CTT ATC GAG CAG CTT CAG GGT      1632
Ala Ser Gly Ser Arg Met Ala Trp Thr Leu Ile Glu Gln Leu Gln Gly
    530                 535                 540

GGC AGC TAC AAG AAG ATT GGC TAC TAT GAC AGC ACC AAG GAT GAT CTT      1680
Gly Ser Tyr Lys Lys Ile Gly Tyr Tyr Asp Ser Thr Lys Asp Asp Leu
545                 550                 555                 560

TCC TGG TCC AAA ACA GAT AAA TGG ATT GTT ATA TCC AGA ACT CAC AGC      1728
Ser Trp Ser Lys Thr Asp Lys Trp Ile Val Ile Ser Arg Thr His Ser
                565                 570                 575

CCA ACC TGA                                                          1737
Pro Thr  *
```

(2) INFORMATION FOR SEQ ID NO: 55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 578 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

Met Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly

```
1               5                  10                  15

Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
            20                  25                  30

Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
        35                  40                  45

Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu
    50                  55                  60

Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys
65                  70                  75                  80

Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys
                85                  90                  95

Val Arg Ile Cys Ser Lys Ser Tyr Leu Thr Leu Glu Asn Gly Lys Val
            100                 105                 110

Phe Leu Thr Gly Gly Asp Leu Pro Ala Leu Asp Gly Ala Arg Val Asp
        115                 120                 125

Phe Arg Cys Asp Pro Asp Phe His Leu Val Gly Ser Ser Arg Ser Ile
    130                 135                 140

Cys Ser Gln Gly Gln Trp Ser Thr Pro Lys Pro His Cys Gln Val Asn
145                 150                 155                 160

Arg Thr Pro His Ser Glu Arg Arg Ala Val Tyr Ile Gly Ala Leu Phe
                165                 170                 175

Pro Met Ser Gly Gly Trp Pro Gly Gly Gln Ala Cys Gln Pro Ala Val
            180                 185                 190

Glu Met Ala Leu Glu Asp Val Asn Ser Arg Arg Asp Ile Leu Pro Asp
        195                 200                 205

Tyr Glu Leu Lys Leu Ile His His Asp Ser Lys Cys Asp Pro Gly Gln
    210                 215                 220

Ala Thr Lys Tyr Leu Tyr Glu Leu Leu Tyr Asn Asp Pro Ile Lys Ile
225                 230                 235                 240
```

Ile Leu Met Pro Gly Cys Ser Ser Val Ser Thr Leu Val Ala Glu Ala
245 250 255

Ala Arg Met Trp Asn Leu Ile Val Leu Ser Tyr Gly Ser Ser Ser Pro
260 265 270

Ala Leu Ser Asn Arg Gln Arg Phe Pro Thr Phe Phe Arg Thr His Pro
275 280 285

Ser Ala Thr Leu His Asn Pro Thr Arg Val Lys Leu Phe Glu Lys Trp
290 295 300

Gly Trp Lys Lys Ile Ala Thr Ile Gln Gln Thr Thr Glu Val Phe Thr
305 310 315 320

Ser Thr Leu Asp Asp Leu Glu Glu Arg Val Lys Glu Ala Gly Ile Glu
325 330 335

Ile Thr Phe Arg Gln Ser Phe Phe Ser Asp Pro Ala Val Pro Val Lys
340 345 350

Asn Leu Lys Arg Gln Asp Ala Arg Ile Ile Val Gly Leu Phe Tyr Glu
355 360 365

Thr Glu Ala Arg Lys Val Phe Cys Glu Val Tyr Lys Glu Arg Leu Phe
370 375 380

Gly Lys Lys Tyr Val Trp Phe Leu Ile Gly Trp Tyr Ala Asp Asn Trp
385 390 395 400

Phe Lys Ile Tyr Asp Pro Ser Ile Asn Cys Thr Val Asp Glu Met Thr
405 410 415

Glu Ala Val Glu Gly His Ile Thr Thr Glu Ile Val Met Leu Asn Pro
420 425 430

Ala Asn Thr Arg Ser Ile Ser Asn Met Thr Ser Gln Glu Phe Val Glu
435 440 445

Lys Leu Thr Lys Arg Leu Lys Arg His Pro Glu Glu Thr Gly Gly Phe
450 455 460

```
Gln Glu Ala Pro Leu Ala Tyr Asp Ala Ile Trp Ala Leu Ala Leu Ala
465                 470                 475                 480

Leu Asn Lys Thr Ser Gly Gly Gly Gly Arg Ser Gly Val Arg Leu Glu
                485                 490                 495

Asp Phe Asn Tyr Asn Asn Gln Thr Ile Thr Asp Gln Ile Tyr Arg Ala
            500                 505                 510

Met Asn Ser Ser Ser Phe Glu Gly Val Ser Gly His Val Val Phe Asp
        515                 520                 525

Ala Ser Gly Ser Arg Met Ala Trp Thr Leu Ile Glu Gln Leu Gln Gly
    530                 535                 540

Gly Ser Tyr Lys Lys Ile Gly Tyr Tyr Asp Ser Thr Lys Asp Asp Leu
545                 550                 555                 560

Ser Trp Ser Lys Thr Asp Lys Trp Ile Val Ile Ser Arg Thr His Ser
                565                 570                 575

Pro Thr
```

(2) INFORMATION FOR SEQ ID NO: 56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1386 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:1..1386

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 56:

```
ATG GGG CCC GGG GCC CCT TTT GCC CGG GTG GGG TGG CCA CTG CCG CTT        48
Met Gly Pro Gly Ala Pro Phe Ala Arg Val Gly Trp Pro Leu Pro Leu
 1               5                  10                  15

CTG GTT GTG ATG GCG GCA GGG GTG GCT CCG GTG TGG GCC TCC CAC TCC        96
Leu Val Val Met Ala Ala Gly Val Ala Pro Val Trp Ala Ser His Ser
            20                  25                  30

CCC CAT CTC CCG CGG CCT CAC TCG CGG GTC CCC CCG CAC CCC TCC TCA       144
Pro His Leu Pro Arg Pro His Ser Arg Val Pro Pro His Pro Ser Ser
        35                  40                  45

GAA CGG CGC GCA GTG TAC ATC GGG GCA CTG TTT CCC ATG AGC GGG GGC       192
Glu Arg Arg Ala Val Tyr Ile Gly Ala Leu Phe Pro Met Ser Gly Gly
    50                  55                  60

TGG CCA GGG GGC CAG GCC TGC CAG CCC GCG GTG GAG ATG GCG CTG GAG       240
Trp Pro Gly Gly Gln Ala Cys Gln Pro Ala Val Glu Met Ala Leu Glu
65                  70                  75                  80

GAC GTG AAT AGC CGC AGG GAC ATC CTG CCG GAC TAT GAG CTC AAG CTC       288
Asp Val Asn Ser Arg Arg Asp Ile Leu Pro Asp Tyr Glu Leu Lys Leu
                85                  90                  95

ATC CAC CAC GAC AGC AAG TGT GAT CCA GGC CAA GCC ACC AAG TAC CTA       336
Ile His His Asp Ser Lys Cys Asp Pro Gly Gln Ala Thr Lys Tyr Leu
            100                 105                 110

TAT GAG CTG CTC TAC AAC GAC CCT ATC AAG ATC ATC CTT ATG CCT GGC       384
Tyr Glu Leu Leu Tyr Asn Asp Pro Ile Lys Ile Ile Leu Met Pro Gly
        115                 120                 125

TGC AGC TCT GTC TCC ACG CTG GTG GCT GAG GCT GCT AGG ATG TGG AAC       432
Cys Ser Ser Val Ser Thr Leu Val Ala Glu Ala Ala Arg Met Trp Asn
    130                 135                 140

CTC ATT GTG CTT TCC TAT GGC TCC AGC TCA CCA GCC CTG TCA AAC CGG       480
Leu Ile Val Leu Ser Tyr Gly Ser Ser Ser Pro Ala Leu Ser Asn Arg
```

145 150 155 160

CAG CGT TTC CCC ACT TTC TTC CGA ACG CAC CCA TCA GCC ACA CTC CAC 528
Gln Arg Phe Pro Thr Phe Phe Arg Thr His Pro Ser Ala Thr Leu His
165 170 175

AAC CCT ACC CGC GTG AAA CTC TTT GAA AAG TGG GGC TGG AAG AAG ATT 576
Asn Pro Thr Arg Val Lys Leu Phe Glu Lys Trp Gly Trp Lys Lys Ile
180 185 190

GCT ACC ATC CAG CAG ACC ACT GAG GTC TTC ACT TCG ACT CTG GAC GAC 624
Ala Thr Ile Gln Gln Thr Thr Glu Val Phe Thr Ser Thr Leu Asp Asp
195 200 205

CTG GAG GAA CGA GTG AAG GAG GCT GGA ATT GAG ATT ACT TTC CGC CAG 672
Leu Glu Glu Arg Val Lys Glu Ala Gly Ile Glu Ile Thr Phe Arg Gln
210 215 220

AGT TTC TTC TCA GAT CCA GCT GTG CCC GTC AAA AAC CTG AAG CGC CAG 720
Ser Phe Phe Ser Asp Pro Ala Val Pro Val Lys Asn Leu Lys Arg Gln
225 230 235 240

GAT GCC CGA ATC ATC GTG GGA CTT TTC TAT GAG ACT GAA GCC CGG AAA 768
Asp Ala Arg Ile Ile Val Gly Leu Phe Tyr Glu Thr Glu Ala Arg Lys
245 250 255

GTT TTT TGT GAG GTG TAC AAG GAG CGT CTC TTT GGG AAG AAG TAC GTC 816
Val Phe Cys Glu Val Tyr Lys Glu Arg Leu Phe Gly Lys Lys Tyr Val
260 265 270

TGG TTC CTC ATT GGG TGG TAT GCT GAC AAT TGG TTC AAG ATC TAC GAC 864
Trp Phe Leu Ile Gly Trp Tyr Ala Asp Asn Trp Phe Lys Ile Tyr Asp
275 280 285

CCT TCT ATC AAC TGC ACA GTG GAT GAG ATG ACT GAG GCG GTG GAG GGC 912
Pro Ser Ile Asn Cys Thr Val Asp Glu Met Thr Glu Ala Val Glu Gly
290 295 300

CAC ATC ACA ACT GAG ATT GTC ATG CTG AAT CCT GCC AAT ACC CGC AGC 960
His Ile Thr Thr Glu Ile Val Met Leu Asn Pro Ala Asn Thr Arg Ser
305 310 315 320

ATT TCC AAC ATG ACA TCC CAG GAA TTT GTG GAG AAA CTA ACC AAG CGA 1008

```
Ile Ser Asn Met Thr Ser Gln Glu Phe Val Glu Lys Leu Thr Lys Arg
                325                 330                 335

CTG AAA AGA CAC CCT GAG GAG ACA GGA GGC TTC CAG GAG GCA CCG CTG      1056
Leu Lys Arg His Pro Glu Glu Thr Gly Gly Phe Gln Glu Ala Pro Leu
            340                 345                 350

GCC TAT GAT GCC ATC TGG GCC TTG GCA CTG GCC CTG AAC AAG ACA TCT      1104
Ala Tyr Asp Ala Ile Trp Ala Leu Ala Leu Ala Leu Asn Lys Thr Ser
        355                 360                 365

GGA GGA GGC GGC CGT TCT GGT GTG CGC CTG GAG GAC TTC AAC TAC AAC      1152
Gly Gly Gly Gly Arg Ser Gly Val Arg Leu Glu Asp Phe Asn Tyr Asn
    370                 375                 380

AAC CAG ACC ATT ACC GAC CAA ATC TAC CGG GCA ATG AAC TCT TCG TCC      1200
Asn Gln Thr Ile Thr Asp Gln Ile Tyr Arg Ala Met Asn Ser Ser Ser
385                 390                 395                 400

TTT GAG GGT GTC TCT GGC CAT GTG GTG TTT GAT GCC AGC GGC TCT CGG      1248
Phe Glu Gly Val Ser Gly His Val Val Phe Asp Ala Ser Gly Ser Arg
                405                 410                 415

ATG GCA TGG ACG CTT ATC GAG CAG CTT CAG GGT GGC AGC TAC AAG AAG      1296
Met Ala Trp Thr Leu Ile Glu Gln Leu Gln Gly Gly Ser Tyr Lys Lys
            420                 425                 430

ATT GGC TAC TAT GAC AGC ACC AAG GAT GAT CTT TCC TGG TCC AAA ACA      1344
Ile Gly Tyr Tyr Asp Ser Thr Lys Asp Asp Leu Ser Trp Ser Lys Thr
        435                 440                 445

GAT AAA TGG ATT GTT ATA TCC AGA ACT CAC AGC CCA ACC TGA              1386
Asp Lys Trp Ile Val Ile Ser Arg Thr His Ser Pro Thr  *
    450                 455                 460
```

(2) INFORMATION FOR SEQ ID NO: 57:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 461 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

```
Met Gly Pro Gly Ala Pro Phe Ala Arg Val Gly Trp Pro Leu Pro Leu
 1               5                  10                  15

Leu Val Val Met Ala Ala Gly Val Ala Pro Val Trp Ala Ser His Ser
            20                  25                  30

Pro His Leu Pro Arg Pro His Ser Arg Val Pro Pro His Pro Ser Ser
        35                  40                  45

Glu Arg Arg Ala Val Tyr Ile Gly Ala Leu Phe Pro Met Ser Gly Gly
    50                  55                  60

Trp Pro Gly Gly Gln Ala Cys Gln Pro Ala Val Glu Met Ala Leu Glu
65                  70                  75                  80

Asp Val Asn Ser Arg Arg Asp Ile Leu Pro Asp Tyr Glu Leu Lys Leu
                85                  90                  95

Ile His His Asp Ser Lys Cys Asp Pro Gly Gln Ala Thr Lys Tyr Leu
            100                 105                 110

Tyr Glu Leu Leu Tyr Asn Asp Pro Ile Lys Ile Ile Leu Met Pro Gly
        115                 120                 125

Cys Ser Ser Val Ser Thr Leu Val Ala Glu Ala Ala Arg Met Trp Asn
    130                 135                 140

Leu Ile Val Leu Ser Tyr Gly Ser Ser Ser Pro Ala Leu Ser Asn Arg
145                 150                 155                 160

Gln Arg Phe Pro Thr Phe Phe Arg Thr His Pro Ser Ala Thr Leu His
                165                 170                 175

Asn Pro Thr Arg Val Lys Leu Phe Glu Lys Trp Gly Trp Lys Lys Ile
            180                 185                 190

Ala Thr Ile Gln Gln Thr Thr Glu Val Phe Thr Ser Thr Leu Asp Asp
        195                 200                 205

Leu Glu Glu Arg Val Lys Glu Ala Gly Ile Glu Ile Thr Phe Arg Gln
```

210 215 220

Ser Phe Phe Ser Asp Pro Ala Val Pro Val Lys Asn Leu Lys Arg Gln
225 230 235 240

Asp Ala Arg Ile Ile Val Gly Leu Phe Tyr Glu Thr Glu Ala Arg Lys
245 250 255

Val Phe Cys Glu Val Tyr Lys Glu Arg Leu Phe Gly Lys Lys Tyr Val
260 265 270

Trp Phe Leu Ile Gly Trp Tyr Ala Asp Asn Trp Phe Lys Ile Tyr Asp
275 280 285

Pro Ser Ile Asn Cys Thr Val Asp Glu Met Thr Glu Ala Val Glu Gly
290 295 300

His Ile Thr Thr Glu Ile Val Met Leu Asn Pro Ala Asn Thr Arg Ser
305 310 315 320

Ile Ser Asn Met Thr Ser Gln Glu Phe Val Glu Lys Leu Thr Lys Arg
325 330 335

Leu Lys Arg His Pro Glu Glu Thr Gly Gly Phe Gln Glu Ala Pro Leu
340 345 350

Ala Tyr Asp Ala Ile Trp Ala Leu Ala Leu Ala Leu Asn Lys Thr Ser
355 360 365

Gly Gly Gly Gly Arg Ser Gly Val Arg Leu Glu Asp Phe Asn Tyr Asn
370 375 380

Asn Gln Thr Ile Thr Asp Gln Ile Tyr Arg Ala Met Asn Ser Ser Ser
385 390 395 400

Phe Glu Gly Val Ser Gly His Val Val Phe Asp Ala Ser Gly Ser Arg
405 410 415

Met Ala Trp Thr Leu Ile Glu Gln Leu Gln Gly Gly Ser Tyr Lys Lys
420 425 430

Ile Gly Tyr Tyr Asp Ser Thr Lys Asp Asp Leu Ser Trp Ser Lys Thr
435 440 445

Asp Lys Trp Ile Val Ile Ser Arg Thr His Ser Pro Thr
450 455 460

(2) INFORMATION FOR SEQ ID NO: 58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1746 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Canis familiaris

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..1746

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

ATG CTG CTG CTG CTC CTG CCG CTG GCG CTG GCG CCG CTC TTC CTC CGC 48
Met Leu Leu Leu Leu Leu Pro Leu Ala Leu Ala Pro Leu Phe Leu Arg
1 5 10 15

CCC CCG GGC GCG GGC GGG GCA CAG ACC CCC AAC GCC ACC TCG GAA GGT 96
Pro Pro Gly Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly
20 25 30

TGC CAG ATC ATA CAC CCG CCT TGG GAA GGG GGT ATC AGG TAC AGG GGC 144
Cys Gln Ile Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly
35 40 45

CTG ACT CGT GAC CAG GTG AAG GCT ATC AAC TTC CTG CCG GTG GAC TAT 192
Leu Thr Arg Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr
50 55 60

```
GAG ATT GAG TAT GTG TGC CGG GGA GAG CGA GAG GTG GTG GGG CCC AAG      240
Glu Ile Glu Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys
 65                  70                  75                  80

GTC CGA AAG TGC CTG GCC AAT GGC TCC TGG ACA GAT ATG GAC ACA CCC      288
Val Arg Lys Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro
                85                  90                  95

AGC CGC TGT GTC CGA ATC TGT TCC AAG TCA TAT TTG GCC CTG GAA AAT      336
Ser Arg Cys Val Arg Ile Cys Ser Lys Ser Tyr Leu Ala Leu Glu Asn
            100                 105                 110

GGG AAG GTC TTC CTG ACG GGT GGG GAC CTC CCC GCT CTG GAT GGA GCC      384
Gly Lys Val Phe Leu Thr Gly Gly Asp Leu Pro Ala Leu Asp Gly Ala
        115                 120                 125

CGG GTG GAT TTC CGG TGT GAC CCT GAC TTC CAT CTT GTG GGC AGC TCC      432
Arg Val Asp Phe Arg Cys Asp Pro Asp Phe His Leu Val Gly Ser Ser
    130                 135                 140

CGG AGT ATC TGT AGT CAG GGC CAG TGG AGC ACT CCC AAG CCC CAC TGC      480
Arg Ser Ile Cys Ser Gln Gly Gln Trp Ser Thr Pro Lys Pro His Cys
145                 150                 155                 160

CAG GTG AGC CGA ACG CCG CAC TCA GAG CGG CGA GCG GTG TAC ATC GGG      528
Gln Val Ser Arg Thr Pro His Ser Glu Arg Arg Ala Val Tyr Ile Gly
                165                 170                 175

GCG CTG TTT CCC ATG AGC GGG GGC TGG CCG GGG GGC CAG GCC TGC CAG      576
Ala Leu Phe Pro Met Ser Gly Gly Trp Pro Gly Gly Gln Ala Cys Gln
            180                 185                 190

CCC GCG GTG GAG ATG GCG CTG GAG GAC GTG AAT AGC CGC AGG GAC ATC      624
Pro Ala Val Glu Met Ala Leu Glu Asp Val Asn Ser Arg Arg Asp Ile
        195                 200                 205

CTG CCG GAC TAC GAG CTC AAG CTC ATC CAC CAC GAC AGC AAG TGT GAC      672
Leu Pro Asp Tyr Glu Leu Lys Leu Ile His His Asp Ser Lys Cys Asp
    210                 215                 220

CCA GGC CAA GCT ACC AAG TAC CTG TAT GAA CTG CTC TAC AAC GAC CCC      720
Pro Gly Gln Ala Thr Lys Tyr Leu Tyr Glu Leu Leu Tyr Asn Asp Pro
```

```
225                 230                 235                 240

ATC AAG ATC ATC CTC ATG CCT GGC TGC AGC TCT GTC TCC ACG CTT GTG      768
Ile Lys Ile Ile Leu Met Pro Gly Cys Ser Ser Val Ser Thr Leu Val
                245                 250                 255

GCT GAG GCT GCC AGG ATG TGG AAC CTC ATT GTG CTC TCC TAT GGT TCC      816
Ala Glu Ala Ala Arg Met Trp Asn Leu Ile Val Leu Ser Tyr Gly Ser
            260                 265                 270

AGC TCA CCA GCT CTG TCC AAC CGG CAG CGC TTT CCT ACC TTC TTC CGA      864
Ser Ser Pro Ala Leu Ser Asn Arg Gln Arg Phe Pro Thr Phe Phe Arg
        275                 280                 285

ACT CAT CCC TCG GCC ACG CTC CAC AAC CCT ACG CGA GTG AAG CTC TTT      912
Thr His Pro Ser Ala Thr Leu His Asn Pro Thr Arg Val Lys Leu Phe
    290                 295                 300

GAG AAG TGG GGC TGG AGG AAG ATT GCC ACC ATC CAG CAG ACC ACC GAG      960
Glu Lys Trp Gly Trp Arg Lys Ile Ala Thr Ile Gln Gln Thr Thr Glu
305                 310                 315                 320

GTG TTC ACA TCG ACT CTG GAC GAC CTA GAG GAA CGA GTG AAG GAG GCT      1008
Val Phe Thr Ser Thr Leu Asp Asp Leu Glu Glu Arg Val Lys Glu Ala
                325                 330                 335

GGG ATT GAG ATT ACT TTC CGC CAG AGC TTC TTC TCA GAT CCT GCC GTG      1056
Gly Ile Glu Ile Thr Phe Arg Gln Ser Phe Phe Ser Asp Pro Ala Val
            340                 345                 350

CCT GTC AAG AAC CTC AAG CGC CAG GAT GCC CGA ATC ATC GTG GGA CTT      1104
Pro Val Lys Asn Leu Lys Arg Gln Asp Ala Arg Ile Ile Val Gly Leu
        355                 360                 365

TTC TAT GAG ACT GAA GCC CGG AAA GTG TTC TGT GAG GTA TAC AAG GAG      1152
Phe Tyr Glu Thr Glu Ala Arg Lys Val Phe Cys Glu Val Tyr Lys Glu
    370                 375                 380

CGG CTC TTT GGG AAG AAG TAT GTG TGG TTC CTC ATT GGG TGG TAT GCT      1200
Arg Leu Phe Gly Lys Lys Tyr Val Trp Phe Leu Ile Gly Trp Tyr Ala
385                 390                 395                 400

GAC AAT TGG TTC AAG ACC TAC GAC CCC TCC ATC AAC TGC ACA GTG GAT      1248
```

```
Asp Asn Trp Phe Lys Thr Tyr Asp Pro Ser Ile Asn Cys Thr Val Asp
                405                 410                 415

GAG ATG ACC GAG GCT GTG GAA GGC CAC ATC ACC ACT GAG ATT GTC ATG      1296
Glu Met Thr Glu Ala Val Glu Gly His Ile Thr Thr Glu Ile Val Met
            420                 425                 430

CTG AAC CCA GCC AAC ACC CGC AGC ATC TCC AAC ATG ACA TCC CAG GAG      1344
Leu Asn Pro Ala Asn Thr Arg Ser Ile Ser Asn Met Thr Ser Gln Glu
        435                 440                 445

TTT GTG GAG AAA CTG ACC AAG AGA CTC AAG AGA CAC CCT GAG GAG ACA      1392
Phe Val Glu Lys Leu Thr Lys Arg Leu Lys Arg His Pro Glu Glu Thr
    450                 455                 460

GGC GGC TTC CAG GAG GCA CCG CTG GCC TAT GAT GCC ATC TGG GCC TTG      1440
Gly Gly Phe Gln Glu Ala Pro Leu Ala Tyr Asp Ala Ile Trp Ala Leu
465                 470                 475                 480

GCA TTG GCC CTG AAC AAG ACA TCT GGA GGG AGC GGC CGT TCG GGG GTG      1488
Ala Leu Ala Leu Asn Lys Thr Ser Gly Gly Ser Gly Arg Ser Gly Val
                485                 490                 495

CGC CTG GAA GAC TTC AAC TAC AAC AAC CAG ACG ATC ACA GAC CAA ATC      1536
Arg Leu Glu Asp Phe Asn Tyr Asn Asn Gln Thr Ile Thr Asp Gln Ile
            500                 505                 510

TAC CGC GCA ATG AAC TCC TCG TCC TTT GAG GGT GTC TCT GGC CAC GTG      1584
Tyr Arg Ala Met Asn Ser Ser Ser Phe Glu Gly Val Ser Gly His Val
        515                 520                 525

GTG TTT GAT GCC AGC GGC TCA CGG ATG GCC TGG ACT CTG ATT GAG CAG      1632
Val Phe Asp Ala Ser Gly Ser Arg Met Ala Trp Thr Leu Ile Glu Gln
    530                 535                 540

CTG CAG GGT GGC AGC TAC AAG AAG ATC GGC TAC TAT GAC AGC ACC AAG      1680
Leu Gln Gly Gly Ser Tyr Lys Lys Ile Gly Tyr Tyr Asp Ser Thr Lys
545                 550                 555                 560

GAT GAC CTT TCC TGG TCT AAA ACG GAC AAA TGG ATT GTT ACA TCC AGA      1728
Asp Asp Leu Ser Trp Ser Lys Thr Asp Lys Trp Ile Val Thr Ser Arg
                565                 570                 575
```

ACT CCC AGC CCA ACT TGA 1746
Thr Pro Ser Pro Thr *
580

(2) INFORMATION FOR SEQ ID NO: 59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 581 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

Met Leu Leu Leu Leu Leu Pro Leu Ala Leu Ala Pro Leu Phe Leu Arg
1 5 10 15

Pro Pro Gly Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly
20 25 30

Cys Gln Ile Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly
35 40 45

Leu Thr Arg Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr
50 55 60

Glu Ile Glu Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys
65 70 75 80

Val Arg Lys Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro
85 90 95

Ser Arg Cys Val Arg Ile Cys Ser Lys Ser Tyr Leu Ala Leu Glu Asn
100 105 110

Gly Lys Val Phe Leu Thr Gly Gly Asp Leu Pro Ala Leu Asp Gly Ala
115 120 125

Arg Val Asp Phe Arg Cys Asp Pro Asp Phe His Leu Val Gly Ser Ser
130 135 140

Arg Ser Ile Cys Ser Gln Gly Gln Trp Ser Thr Pro Lys Pro His Cys
145 150 155 160

Gln Val Ser Arg Thr Pro His Ser Glu Arg Arg Ala Val Tyr Ile Gly
165 170 175

Ala Leu Phe Pro Met Ser Gly Gly Trp Pro Gly Gly Gln Ala Cys Gln
180 185 190

Pro Ala Val Glu Met Ala Leu Glu Asp Val Asn Ser Arg Arg Asp Ile
195 200 205

Leu Pro Asp Tyr Glu Leu Lys Leu Ile His His Asp Ser Lys Cys Asp
210 215 220

Pro Gly Gln Ala Thr Lys Tyr Leu Tyr Glu Leu Leu Tyr Asn Asp Pro
225 230 235 240

Ile Lys Ile Ile Leu Met Pro Gly Cys Ser Ser Val Ser Thr Leu Val
245 250 255

Ala Glu Ala Ala Arg Met Trp Asn Leu Ile Val Leu Ser Tyr Gly Ser
260 265 270

Ser Ser Pro Ala Leu Ser Asn Arg Gln Arg Phe Pro Thr Phe Phe Arg
275 280 285

Thr His Pro Ser Ala Thr Leu His Asn Pro Thr Arg Val Lys Leu Phe
290 295 300

Glu Lys Trp Gly Trp Arg Lys Ile Ala Thr Ile Gln Gln Thr Thr Glu
305 310 315 320

Val Phe Thr Ser Thr Leu Asp Asp Leu Glu Glu Arg Val Lys Glu Ala
325 330 335

Gly Ile Glu Ile Thr Phe Arg Gln Ser Phe Phe Ser Asp Pro Ala Val
340 345 350

Pro Val Lys Asn Leu Lys Arg Gln Asp Ala Arg Ile Ile Val Gly Leu
355 360 365

```
Phe Tyr Glu Thr Glu Ala Arg Lys Val Phe Cys Glu Val Tyr Lys Glu
    370                 375                 380

Arg Leu Phe Gly Lys Lys Tyr Val Trp Phe Leu Ile Gly Trp Tyr Ala
385                 390                 395                 400

Asp Asn Trp Phe Lys Thr Tyr Asp Pro Ser Ile Asn Cys Thr Val Asp
                405                 410                 415

Glu Met Thr Glu Ala Val Glu Gly His Ile Thr Thr Glu Ile Val Met
            420                 425                 430

Leu Asn Pro Ala Asn Thr Arg Ser Ile Ser Asn Met Thr Ser Gln Glu
        435                 440                 445

Phe Val Glu Lys Leu Thr Lys Arg Leu Lys Arg His Pro Glu Glu Thr
    450                 455                 460

Gly Gly Phe Gln Glu Ala Pro Leu Ala Tyr Asp Ala Ile Trp Ala Leu
465                 470                 475                 480

Ala Leu Ala Leu Asn Lys Thr Ser Gly Gly Ser Gly Arg Ser Gly Val
                485                 490                 495

Arg Leu Glu Asp Phe Asn Tyr Asn Asn Gln Thr Ile Thr Asp Gln Ile
            500                 505                 510

Tyr Arg Ala Met Asn Ser Ser Ser Phe Glu Gly Val Ser Gly His Val
        515                 520                 525

Val Phe Asp Ala Ser Gly Ser Arg Met Ala Trp Thr Leu Ile Glu Gln
    530                 535                 540

Leu Gln Gly Gly Ser Tyr Lys Lys Ile Gly Tyr Tyr Asp Ser Thr Lys
545                 550                 555                 560

Asp Asp Leu Ser Trp Ser Lys Thr Asp Lys Trp Ile Val Thr Ser Arg
                565                 570                 575

Thr Pro Ser Pro Thr
            580
```

((2) INFORMATION FOR SEQ ID NO: 60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16862 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:3415..3440

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION:3441..3903

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:3904..3988

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION:3989..4689

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:4690..4893

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION:4894..5645

(ix) FEATURE:
(A) NAME/KEY: exon

(B) LOCATION:5646..5831

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION:5832..7181

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:7182..7202

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION:7203..8307

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:8308..8803

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION:8804..12266

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:12267..12401

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION:12402..12815

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:12816..12986

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION:12987..14085

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:14086..14187

(ix) FEATURE:

(A) NAME/KEY: intron
(B) LOCATION:14188..14473

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:14474..14539

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION:14540..14998

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:14999..15190

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION:15191..16862

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

GATCATATTA ATTTGAAGGT GGCGGGGCAG GATGGTTCTG TGGTGCAGTT TAAGATTAAG 60

AGGCATACAC CACTTAGTAA ACTAATGAAA GCCTATTGTG AACGACAGGG ATTGTCAATG 120

AGGCAGATCA GATTCCGATT CGACGGGCAA CCAATGAAAC AGACACACCT GCACAGTTGG 180

AAATGGAGGA TGAAGATACA ATTGATGTGT TCCAACAGCA GACGGGAGGT GTCTACTGAA 240

AAGGGAACCT GCTTCTTTAC TCCAGAACTC TGTTCTTTAA AGACCAAGAT TACATTCTCA 300

ATTAGAAAAC TGCAATTTGC TTCCACCACA TCCTGACTAC TACCGTATAG TTTTCTCTAT 360

TCTTTCATTT CCCCCTTCCC CATTCCTTTA CTGTACATAA AGTAACTGGT ATATGTGCAC 420

AAGCATATTA CTTTTTTTTT TTAAAACTAA ACAGCCAATG GTATGTTTTG ATTGACATCA 480

AGTGGAGACG GGGGGGAAAA TACTGATTCT GTGAAAATAC CCCCTTTCTC CATTAGTGGC 540

ATGCTCATTC AGCTCTTATC TTTATATTCC AGTAAGTTAT TTTGCTCTCA CTGTTTTAAC 600

AACAACAACA AAAAAACAAC AACATAAAAA TCCTTGCATA CCTTGTTCAA TTGGAGAATT 660

TTAATGTTTT TCATTTATCA TTGTAAAACC AAGGACAATT TTATAACTTT TTTGTACTTA 720

GCTGTTACAT GCAGAGCAAT CTGTCTTTAA GTAGGGATAA ATTACTCTAA AACAAAAAAG 780

AATCCTAGAT AGTTTTCCCT TCAAGTCAAG CGTCTTGTTG TTTAAATAAA CTTCTTGTTT 840

AAAAAAAAAA AAAGTAAAAA AGAAAAGTTA TGCAACAATT AATGGCCCAG AGGCAATCCT 900

TGTTAACATT TTGATGCATC TTTTAGCTGT TTTTTTTTTT TTTTTTTTTT TTGACTGAGT 960

TTGACTCTTG TCACCCAGGC TGAAGTGCAA TGGCATGGCA TGATCTTGGC TCACTGCAAC 1020

CTCCGCCTCC CGGGTTCAAG TGATTCTCCT GCCTCAGCCT CCTGAGTAGC TAGGATTACG 1080

GGCATGCACC ACCATGCCTG GCTAATTTTG TATTTTTAGT AGAGTTGGGG CTTCTCCACA 1140

CTGGTCAGGC TGGTCTCGAA CTCCCAACCT CAGGTGATAA GGGAAGGGGC ACTATTGACA 1200

TTTATGGTTG GGGCAGAGGT GTAAGATATT CTTCAAAGCA CTACCTACAT GTTGAAGAAT 1260

TGTTCCTCAC CCAGATTCTC AAAAGTCCCC CAGGACATTC ACGTAGTGAA AACCTGTGTT 1320

TAATTATCTG AGCCTATAAC TTAATACAGT TTTAAAATTT TTTTTTAAAT ATACAGTGAA 1380

CTTTCTAGGA ATGCAATTAT AGTTGTGTGT AAAATTAGGG AAAATTAACT TTGCTACCAA 1440

GAGTTGTTCA ACATTTTGTT AAATCACTTC ATTGATGGCA ACATGCTGGA GGTAGTTGAG 1500

TCACCAACTC AGCACCTGGA TCAGCCTGTG TTGGTAGCAG TTTCATCCCC GTGGTTCTGT 1560

GAATAGGTGG AAGCATCTGC TTACTCCATC AGGACTTCTA GGGTAGTCGG GCCTTGGCAC 1620

TCACACATTA AAATACTGTT TATGTTATTT TATTGCAAGT TACTTTTCTT TCATTTCCCC 1680

TTTACGTTAC AGAAAGGGAA GCATTTTGCT TTCTGTTTAA AGTTGTGTAT GTAGGTAGGT 1740

TATATCATCT AWGACTTTCT CTCCCTCCTT CCCTTTCTTT TTGTTTGAGA TGGAGTCTTG 1800

CTCTGTCACC CAGGCTGGAG TGCAGTGGTG CGATCTTGGC TCACTGCAAC CTCTGCCTCC 1860

CGGGTTCAAG CGATTCTGGT GTCTCAGCTG GGATTACAGG CGCACACCAT CACACCACGC 1920

TAATTTTTCT ATTTTTAGTA GAGATGGGGT TTCGCCATGC TGGCCAGGCC AGGCTGGTCT 1980

CAAACTCCTG AGCTCAAGTG ATCAGTCCGC CTCGGCCTCC CAAAGTTCTG GGATTTCAGG 2040

CGTGAGCCTC ATCTATGAAT CTCAATTTAG GACAGTAAAA GTGTCATTAC AAAAATATTT 2100

ATTGTAAAAA AGGGTTGGAG GTTGAGAATC TCAATTCTAG TCAGTCTCTC AGTGTTTGGT 2160

TTCTTCCTAC CATTTTTCCC CCTAGGACCA GCCAGAAAGC AGCTTTTTTT TTGTCCCCCC 2220

CAACAAGGAG CCCACTGTTT CCTCTCCCAG CCCAAACTCA GGCCTACGAA CAACAACAGC 2280

ACTACACACA CACACACACA CACACACACA CACACACACA CACCCCTCCA CTTCAAGGTA 2340

TAGCCAAGAG CTTCTGGAGC CGTCAAAAAG GTCTGTACCT GCTGTCTTTA GAGCTTCCAG 2400

TTTGCCCTTG GTCAAGAAAT ACTGTTTGCT AGGCTCTGCT GGAGTACATC AGGTAATACT 2460

GGCTTCTAAA CCACCCTGAG GTTCTTTTCT CTTGTCCTTT TACTCCCTTC GTACTTCAAT 2520

TTCTCTCCTT GATGTCCCCC TCCCTGTTTT GTTTTTTGCC TCCAATCCGT TCTGCGCGTT 2580

CCCTGCAGAG CAGGCGAGTA GCAATGCTGC TGGACCATGG AGCTGCTCTA GTCTCCCAGA 2640

AATCTCTTCT ACACCCAACC CTTCTTGCGC TTAGGTGGTC CTCAGTCCCC CTCCCCCACC 2700

TCCTTCTGAC CCAGGCTTCT TTCTCGCCCT CCGGTCGCAG TTCTCCTGGG CATCTGCCTC 2760

TGCCTCTCTC CTCTCACCCG GATCTAGGGC TGCCTTCTCT TTGTGCAGCC GTCTTTCTCC 2820

ACCTTCATCC CAGACTCCCT GTCTCAGCGC CAGCTCCTCT GCCTTTGGCT CGGGTTCCCT 2880

CTCCCCCACC CCAGCTTCCA GTTGTTTGGC CCGCAGGTCC CTCGGCAGTG ACCGGCGCCC 2940

CCCGACGAGT GCGTGTGCAC CAGGGCACCT CCCTCTCCCC CACCTCTCAG CCCCGCGCGT 3000

CTCCACCGCC CGCCCCACCG CGCTGTGGGC GGTCCAGGGC GGGGCTGGGA TCCGGGGCGG 3060

CTCCCGGGGC TCGGGTTGTG GGAGGCGCCC TCTCCCCGGT CTTCCCCTCT CTTCCCCCCG 3120

CCCTGCCTTC CCTTGCACCC TCCTTCTTCC CTCCGCCCGG GAGCTCTCCC TGGTCCCCGG 3180

CGCCGCCTCC TTCCCTCCCG GCTCCCCGCT CCCCGCTCCC GTGGCTGCCG CCGCCCCGGG 3240

GAAGAAGAGA CAGGGGTGGG GTTTGGGGGA AGCGAGAGAG GAGGGGAGAG ACCCTGGCCA 3300

GGCTGGAGCC TGGATTCGAG GGGAGGAGGG ACGGGAGGAG GAGAAAGGTG GAGGAGAAGG 3360

GAGGGGGGAG CGGGGAGGAG CGGCCGGGCC TGGGGCCTTG AGGCCCGGGG AGAGCCGGGG 3420

AGCCGGGCCC GCGCGCCGAG GTAAGAGCCA GGGCCCCGGG TTAGCAGGGC TCGGAGAGGG 3480

GGCGCGCGGC GTGGTGGGGG AGGGGGCAGT GGGCGCAGGG CCCAGCTGGG GGAAGCGGGG 3540

CTGGGGGAGA GGAGGAACCG CGGGGATGGA ATCGGGGAGC GCTGAGGCGG CCGATGCCGG 3600

GAGCGTGGGT AAGCCAGGCT TCTGCGAGCC GCGGGGGCCG GGGGAGAGGA GGTGGTGAGA 3660

GGTGGAGTCC GGGAGGGTTG GGGGCCGAGG GAGGCAGGAG GAGGGTGGGG ACAGGCTTTC 3720

TCTCCTCCTC TCCCCCCACC CCGCGCGGGG CTCCGCCCCC GCCTCCTCCG CGGGGCGCTC 3780

TCTTGGTCCC CAGGCTGAGC CCGGTCGGAG CCTGCGAGGC AACCGGCAAG AGGTCGAGTA 3840

GTCTCCGGGT GCGGGCCGCG CCGGCGGGGC TCGGTCCAGT CCTCATGGCC GCCTCTCACT 3900

TAGATGTTGC TGCTGCTGCT ACTGGCGCCA CTCTTCCTCC GCCCCCCGGG CGCGGGCGGG 3960

GCGCAGACCC CCAACGCCAC CTCAGAAGGT GCATCCTTCT TCGACGACCT CCGGCCCTCC 4020

TTCGCTCCAC TTCCCTTTCC CTGCATCTCC TCATTTCTGG TCCTCATCAC TATCCCATCA 4080

GTCCCACATA TCATCCCGGT CTGGCAACCC CTTCTGCTCG GCCCGACTTT ACTACTGCTG 4140

ACCTCCTTCT GTCACCCCAC GTTACTATCC AGCACCTCTT TTCTCTGCCC ACATTGCTAC 4200

ACTATACCAC CTTCCTGTGC ATTTTCTCCG CCTCAATCCC CTTTCCCAGC CCCACATTAC 4260

TACCTCAATT ACTCCCTTTT CTTGGTCCCA CTTTGCTGTC CAGATGATCT TATTAGCCTC 4320

CCTTTATCCT CCTATCCTAA TTCAACTGGA ATATCCTCAT TTAGCCTTTT TTTTTAAAGA 4380

AAAGCTCCAC CCACATATCA TACCCTTCAT GATTTCTTAA TTACTTTTCT TTCTTACCTC 4440

CACCCAGCAC CCTTCCCTCC CCACTTGTGG GTTCTCTCAT CAGCTTTAAC CCTGGCCCTT 4500

TACTCTCTGT CCTTTAGCCA GGGGATCTGT ACCTGTCCCC ACTCCCACCC TCTAGTGCCC 4560

CATCCCTCTT CCTCTGTCCC CAGCCTGCCC ACAGACCACG CCCTACTCTC CCCTTCCTCC 4620

CACTGGGGAG CCTGCCTTTT CCTCTTTCCC ACCATTCCTC TCTGTATGCC TCCCCGACTC 4680

ACCCCTTAGG TTGCCAGATC ATACACCCGC CCTGGGAAGG GGGCATCAGG TACCGGGGCC 4740

TGACTCGGGA CCAGGTGAAG GCTATCAACT TCCTGCCAGT GGACTATGAG ATTGAGTATG 4800

TGTGCCGGGG GGAGCGCGAG GTGGTGGGGC CCAAGGTCCG CAAGTGCCTG GCCAACGGCT 4860

CCTGGACAGA TATGGACACA CCCAGCCGCT GTGGTGAGTA GCCTCGGAAG CCCCTCCCCT 4920

CTTCAAGACT ATTCCTTTTC CTGCCGCAAA CTTAGCATTA CTGCTTGCAA GTCAGCACTT 4980

TAAATCCAGT ATACCAAAAT TCACAAATAC ATTTATTGAA TGACTACTAC ATAAGAGCAA 5040

TTTTGCTCTG TGCGGTTGGA GGTAGTAGAG CTAGCAGCCT GCACAGTTCA.TTTCATCCTC 5100

CCTTCATTAG GCCACTGATC ATTGGCCTAT AACATTGATA ATTCATCTTG TCAGTTATTC 5160

TCTTTGAGGA TCATTAGTGG CAGATGATGA CAAAAAAATT CTAAAATGAT TTCATCACAT 5220

TTTTGAATAC CTCTGTCACC AACCCAGAGA CCATATGCCC AAGAAACAAA AGCCAGTTTA 5280

ATATTAATAG AAGCCAACTA TAATAAGAAA AGCAAATCTG ATTGTGCATC CAAAGTTATA 5340

TACATCTACA TATTTCAAAG CCAGAGAACC GCCCACTGTA GCTGACTTTG AAGAGATCCC 5400

ATTTTGTGTG CTTATAGCCC CATCTTGGGT TCCTAAAATG GTAATTTTTT TTTTCTTTTG 5460

GGAATGTGTG GATGCTTGCA CAGGTAAGGG AGGATTGGAA GATAGGTAGG CAAATCCTTT 5520

TCACATGTGA TTTTCTTTAG AGCAGGATGC TTGTGGACCC AAACCTGCAC CTGAGTCCCC 5580

TGCTCTTTAA AGGGAAAGAG CCTTCTTCAA CTCGCCTCTC TTCTTATTTT CCTATCTCTC 5640

CACAGTCCGA ATCTGCTCCA AGTCTTATTT GACCCTGGAA AATGGGAAGG TTTTCCTGAC 5700

GGGTGGGGAC CTCCCAGCTC TGGACGGAGC CCGGGTGGAT TTCCGGTGTG ACCCCGACTT 5760

CCATCTGGTG GGCAGCTCCC GGAGCATCTG TAGTCAGGGC CAGTGGAGCA CCCCCAAGCC 5820

CCACTGCCAG GGTGAGGGGA ACAGCTGCCT GCATGCAGCT GATGAGGACG CTTGTGTGAG 5880

GATGGGAGTG GGGTGGGAAT GGATAATGGG AAAGAATGGA GAGCTATAAA AATGTGGGGG 5940

AGGACACTGG AAAGGGGAGA TGAAAGTCCC TTTTTCCTCC ATCACCTGCC TCAAACTTCC 6000

TCTTGCAGTC CCCGGTATCC TCTGTAGGTT GGGGGCTTCC TTCCTTTACC TTTTAAAAAA 6060

ATCTTCCTGC TCCCGATTCT TAGACCTCAC GTTTTCTCTT TTCCTTTATG AATCTCACCT 6120

CTCTCACCTT CTTCAGGTTT AAATACTCCA ATTTTCCCTT TCTCTAAACT TAGAAATTTC 6180

CATGCATCAC CCTCTTCTAG AATTCATCCC TCACCATTCC TTATATAATT GATTTATTGT 6240

AAAGACTCAG AAATAAATCA AACATTCTAC TAAGAAAAAT TGAGAAGGGG AGCTCTGGGG 6300

GTGGAAACAT ATTAGGGTAA AAGACTTAAA ATTGGAGGCA GCATTATCAG AAGATGAAGA 6360

ACAACTCAGG GATGGGGTGG GAAGAAGACA GGTCCTTTTC TGKACTTCCT AGACAACCTC 6420

CATTATTCCC TAAGGGAATC AGTGTTGTGT CTGTCTACYT TTTTTTTTTT TTTTTTGCCA 6480

CGTAATTTTA CAAACTCTCC CTTTTCTAGG CACCCGAACT CTCTGCCATC TTCTCTCCTG 6540

GGATGCAGTC ATCCCATTTG TATGCCTCAT ACTTCCTCTA CCCTGGTAGA TTCTTTCAAG 6600

ATCCTTGGGC TTTACTTTCC TCACATAACT CAGTTATTCT GCTTCTAGTT TACCATTTTA 6660

TTCTGGAAAT TGAGAGTCCC ATCCAGGGGT GGACTTATGA CACTACTGAA ACTTAGACTT 6720

CAAGGTTCCT CACCTACAGG GCCCTCTTCC TGTGCTCTAA TAATATAGAG GGCTCGATGG 6780

ATATGTGTTC ATATGGTAAC AGGCTTTTGT AAAAATTGCA GAAATAAGAT TTTAACAGCA 6840

ATTGCTTAAA GCCAATTGTA TGTGTAATTT TTTTTCTTAA AGACTCCCAA TTTTGTAATA 6900

TTCAGGCACC ACAGAACCAA GATCTGCCCC AAACTTAGCT ATTGGCATTC CCGTCTCAAA 6960

TTCTGTTGTC CTATGAAAAA TCGAAGAAGA AAATAAGTCC TGACCCCCTT ACCCCCAGAC 7020

CCACCTTGTT CTTATCCCCA GGCACCCTCC CCTCAGAAAC GCAGGCTTCT GCTCTCCCCG 7080

GTCTTCAGCA TGGACAGGTG TGGGAGGGGG CTGGGGATCA GGCCAGGGAA GCTGGGCGCC 7140

AGTGGTAACT CTTCTCTGAT CCCCGTCTTT CCTGCTGCCA GTGAATCGAA CGCCACACTC 7200

AGGTGAGATG AGAAACCCTT ACCGCGCGCA CTGCAATGCC CTCCCCTTCA CTCTGCACCC 7260

TCCACCCCCC TGAAATTCTG CCCTTAGGCT ACGGGGCGTC GTCCTTTCGC ACCTTCCCCA 7320

ACCCACCCCA GTTTGCGGCC ACCCCCTTCC CTCCCTACCT GTTTCCTGCC TCCAGTCCCG 7380

GTTTTCCACG AGGCTGCGGT CTCTCCTTGT CCCTGCTTGG CTACACTTCC CTGGGCTCCA 7440

CCTCCTCCCA GACTGAGCCT CGCCGGTGTC AGGCAGAGCC CAGCAGARGG CGGCAGGGTG 7500

CTGGGAGACC CTGAGCTCCC ACCACGTTTT CCCCTGTGGG GTTCCTTGCG ACCTTCGCTG 7560

GAACCTTTTC CAGCCTGCTG CCTCCTAGGA TTTCACCTAA TGGACTTTCT CAGCCTGTCC 7620

CACCCATCCC AACCCTGGCC AGGCCTCTCG CGCTCTTCCC CACATCTTTT CCTTCCGTGT 7680

ACCCCTTCCC TCGTCTTTTC TCAATTCCAT GTCCTGTCTC CCTTTCTTAG GCTTCTGTCT 7740

ACCCAGCCCC AGGCTCCCTT CCACGACCCC ACCACTCCCT CAAACCAGCC TCCCTTCCGT 7800

ACCCAACTCG TTCCCTCCAA AACCGTTTCC TCTCCCCCAC ATCCTCAGTG CTTCACTGTA 7860

TCGACTCATA CTCCCACTTC AGACCTCAGG CGCCAGCCCC GTTTCTCTCC CGTCCCACTC 7920

GCATCCTTCC CTTCCTACCC TGGTTCCTCC GTGCTTCAGC CTCCCGCGGC TCCCTCCGCC 7980

CACCCCGCCC TCCTGGCACG CCCCGTCCCC ATTTCTCCTC CCCTCGGGTC CCCTTAAGTG 8040

AGATCCCTCC CTTCCTCTTT CGTTCCTTTC CTCCTCGAGG TTGCATCCCC CCTCCCCTCC 8100

CCGCCCCTCC GACTGTCGCT CCCACCTCGG CGCTCGCTTC CCTCCCCGCC CCCTTCCTGC 8160

CTCCCCAGCT CCCGCCCGCC CCCCCACCCC CCGCTGCCGC GCGCCGCCCG TGACGTCAGA 8220

GCCCCCTCCC AGCCCCACAT CTCCCTCCTG CTCCTCCTCC TCCCCTCCGT CGGTCAGTCA 8280

GTCCGCGAGG AGAGTCCGCG GTGGCGGCGA CGGTGGCGAG AGCCGCGGGG GCCGTAGGAA 8340

GCCAACCTTC CCTGCTTCTC CGGGGCCCTC GCCCCCTCCT CCCCACAAAA TCAGGGATGG 8400

AGGCGCCTCC CCGGCACCCT CTTAGCAGCC CTCCCCGGGA AAAGTGTCCC CCCTGAGCTC 8460

CTAACGCTCC CCAACAGCTA CCCCTGCCCC CCACGCCATG GGGCCCGGGG CCCCTTTTGC 8520

CCGGGTGGGG TGGCCACTGC CGCTTCTGGT TGTGATGGCG GCAGGGGTGG CTCCGGTGTG 8580

GGCCTCCCAC TCCCCCCATC TCCCGCGGCC TCACTCGCGG GTCCCCCCGC ACCCCTCCTC 8640

AGAACGGCGC GCAGTGTACA TCGGGGCACT GTTTCCCATG AGCGGGGGCT GGCCAGGGGG 8700

CCAGGCCTGC CAGCCCGCGG TGGAGATGGC GCTGGAGGAC GTGAATAGCC GCAGGGACAT 8760

CCTGCCGGAC TATGAGCTCA AGCTCATCCA CCACGACAGC AAGGTAGCCC TGGACATGGG 8820

GGTGGGTGGG AGGTGGGGGC TTGCGGGSCA GGGGGCCAAG CAAGCCTGCA CGCGCCCCCA 8880

TCTGTCTGAG TCGTCTCTGG GATTGCGAGG CAGACCCCTC CCTTGTGTGA CTGGCAGGAG 8940

ATGGGCTGGG GGTGCAGGAG CTTGGGAAGA GTCGCAGGGG CTGGAGGTCC AAGATGAGGG 9000

TCTAGGGGCT CAAGATGGTT AAGCATGCTG CAAGGCAGAC CCTTCTGCCC CGCTGCGGGA 9060

GTCTCGCAGA AGTGTCGGGG TTTGGAGAAA CTGGTGGTGG ATTTAAGGTA TTAGGAGACA 9120

CTGATCCTCT GAGGGAGTAA ACTAACCCTG GAATGGGTTG GGGGTGGAGG GAATGTCAGA 9180

GGTGGGGAGC TGGATTGGGG GGTTACATTT ACCATGGTAA CAAGGTAAAA TCTTGGCGTA 9240

GGTTGGAGCT GGAAGGAATA GGGACAGAAT GAGGAAAATT TTGAGAGACT TGAGAGCTCT 9300

AGTTTATTTA TCTTAACAAA ACAGCAAGGT AGTGGTGAGC CCTACCTGAC TCCTTCTCAT 9360

CCTTCTATTC CCAACCCTGT TGAGCATTCC CAGACTGTGG GATAGATGGC ATATGGTGAT 9420

TGGGGAAGGC TAATGATCAA GAGGTGGGCA GAGGCACTGG GAAAATGAAT TGGATTGGGG 9480

ATCCACATGG GAACCCCCAC AATAGCATGG GGATGAAGAA GAGTCAACAT ACAAGGAGAA 9540

GAGAACAGAA AAGAATGGCA GTGGGGGAGA GGGGCAAGGA GGTAGCGTGG GGATAATGAG 9600

AGATCTTGGG GCACCTTATG GAACTTGGGT CCTGACCTTC CCTTCCCTTA TAGCATTGTG 9660

GCCTCTAGGA TGTGAGAAGG GAAATGGGAT GTAGGGATTA GGGAGGTGAG TTGAGGGAGA 9720

GAGAGAAGGT AAGCAAATTT GGGTCCAGGG GTATTAGGGG ATAGCTTATA ATGAGGTTTT 9780

TTTTCCCACC CCTCTCCCCT ACATGAATAA TTGGGGGTGC AGGGAAGGAT GTGACACAGG 9840

GAAGGAGATT TAAGATCTCA AATTTATCTT CACTGACATG TGGCCCCAGA GACTTAAGGA 9900

ATTGGGTTAG GGTGAAATAG AGTACACAAG GTGAGAATTT GGTGATCTTA CCAAATATCA 9960

ACCTTGGGGT GATCCAAGGA TTTATATTCA TTTTTAGAAC ATCACTATAC ACCTAGAAAT 10020

AGGTGTGTGT CTGGGATAGG TGTGTGAGGG GACAGAAGTG AGGTTGAAGG TAGGGTGCTT 10080

GAAGAGAAGA GAGCACAAGG ATTATCAGGA GCTTGGCAAG AGAACTTAAA ATCCTTTTTG 10140

ACTGTTACTT TCTCGTGGTT CTCAGCCTTC AGTGTACATA AGAATCACCA GAGGAGTTTG 10200

TTAAAAATAC AGATTCTAGC TCCTTGGTCA GGGATGAATC CCAAGTATTT ATCTGTATTT 10260

TTACTAATAG ACATCCTATC TTGGTGGATT CCTGAGCTGT AAGCTAACCC CAGAATGCCT 10320

ATGGGAAGAG CAGCAGGGTA CAGGAAAATA ATTAGGTATT AGGGTACGGG AGGCAGGAAG 10380

AGAAGTAGAG GATCAGATCT GGTAGAGGGT CAGACTTGGG ACAGTCAGAG AGATCATTGG 10440

TTTTGGGGAG TGGAGTGTGA AGAAAATGAC AGGGAGAGAT GGGTGCAGGC TTTATGATAG 10500

GGGATCACAG GAGATAGGGG AGGCCTGGCT GTGAGCTCAA ACTCATCCAC CATGACAGGT 10560

GATTCCCTGG AGGTGGCGGG GAGCAGACGT GGGACCTGGG AGAAGGGAAC TGGAGAACAT 10620

CAGAGGCATC AAGCGGGGTG GGATGGGAAG GCAGAAGAAC CAGAATGTGT CAATTGGAAT 10680

GAGTCGGTTT CCTGCCTGCA AATCCAGATC CTTGCAAGAG CAAAGAGAGG GAGGAGAACT 10740

AAGGAAATCT ATTGGGGAGG GGGAGAGAAT CACGTGGTGG AGAGAATCTG CAGTGATGAA 10800

TAGTGTGTGG AAGAGGGAAA CGGTTGCAAG AAAAGGTAGA TAAGAAATCA GGAAACAAAA 10860

TGGGGGGCAT GCCTTGCCCT GTTGATATGT ATCTTATATG TTCTTGAATG TCCTCATTGK 10920

TCCTATTAAC CCTGTCTTTA GAGAAGTGGA GGGGCACTGA GGGGCTGTGG GAGAAGCTGG 10980

GAGCAGGATC TGGAGTAATA GATGTGGGGA GAGTGCAGGA AGGTGGGTCC TGAGAATGGT 11040

AAAGATTTAC AAAGTTGCCC TAGTGGGAGG CATAAAGAGA AAACCTTCCA ATGTTGTTGA 11100

GCACTGCCCT TGGCCAGAGT GAGGGTAGGG TGGGCAACAG AGAATTCTCA GTGACTGCTG 11160

GTTCTTCAGA TTCCAACAGC TTCCCCTGGC TCCCCCTTCT CCAACTTCCC ACCGTGTCCC 11220

AAATGTCAGG CCTCAGTGGG AGGTAAGCAG GCTCCAGAGT GCTTTCTTTA TTTCCTTTCT 11280

ACTTATCCTC CCCTCCTGGC AACATTTCAC CCTCCTTAGT CCCCTGAGCC CCCTGTCTGT 11340

GTCCCCTCTG CCCTGGCTCC CCACTGGCTG CCATTTCGTC TTCACATGCA TTGGGGTTCC 11400

AGCAGCTTCT GAAATGTCAT ATATCAGTGG GAGGGGAACA GGCAGTGGGA GACCCAAGGC 11460

TGGCTCTTCC TCCCCCATTT CCCCTCCTCC CAAGCTTCCT TTCTTCTCCA GCTTTCTGCT 11520

TGTTTACTTT CCCTAGCTCC AAGCCTCTCT TTAAGGCACC TCTCAAATTG KCTGGTTTCT 11580

TGAGAGTTCC ATTCTATTCA TTCTCTCTGT TCTTTCCTCA TCCTACATTC TTCCCTACTT 11640

CCACCCCCCA GTGTCTTTTT TTCTAATGGA CCTGTCAAAT GTCAGCGCCC AGCAGGAGGG 11700

ATGGATCACT GAGCGGGACC CCCTACTGGT CTTGTTCCTG TTCTCTCTTT ACTTATCACT 11760

AGCTCTGAAA AGAGAAGAGG GAGGAAACAA ATGGAAGGTG GGGAGAAGGG GTTTGCAGAG 11820

GTGAGGAAGG AATTTTCATA ATATGGCTTT GAGCAAGCTA TCTGGGGATG TGGAAAGAGT 11880

TTACCGTATT CCTACTGACT TCTTCCACCC ACTGGTGTTT GAAGCATAGA AACATGGGGT 11940

AAAGGGCTTG GTGACAGAGG GAAGGGGGAT GTCTGAGGGT GAGCTGAAAG GAGGTAAGGT 12000

GGTATGTTCA TTAATACCAA AGGAGGGGTG TGCAGGAGAG GTGATGGGTA AGGCTCCAGA 12060

TGGAAGACAG AGAAGGAAGT TTAATGAAAG AGGAGAAAAA AGGCACTTGA CAGGAAGAGA 12120

TGCCAGAAAG GAGAAGAAAA CGGTAATTAA TGATGAAAGT GAGTAATTGA GAAAGGAACT 12180

AATTTGTTCG AGAAAGATAA GAGCAGGAAT TGCAGACAGG GGAGGGGCCC CAGGAGAGCT 12240

TGCCCTCATC TCCTCTTGTC TTTCAGTGTG ATCCAGGCCA AGCCACCAAG TACCTATATG 12300

AGCTGCTCTA CAACGACCCT ATCAAGATCA TCCTTATGCC TGGCTGCAGC TCTGTCTCCA 12360

CGCTGGTGGC TGAGGCTGCT AGGATGTGGA ACCTCATTGT GGTAAGCAGG GCTATGGGGG 12420

TCAGAAGATG GGGTCATTCC CTTTTGAGCT CTACTGAAGG GACGATGGCG ATTGTGGGTT 12480

TGTATTGAAA AGGAGTGTGG AGGACCTGCT ACTAAGATTC AGAGTCCTCT GCAGACCTGA 12540

GCTAGGCAGC CTCCTAGCAA CAGTGGCCTG ACAGTGCTGC AGCTGACCTC CTTCTTCAGA 12600

AGGAATTGAA ATTAGATCAG TGAAAGAGCA TCCCGGTTGT GAGGGGTGTG TGGGCCTTTG 12660

AGAATCTCTT TTCCTTAGGC AGACCAGAGG TGGGGAGGTT TGGAGAGAGT AAGGAAGAGA 12720

AACCCAAAGG CAGGAAGAGG GTTAAAGGAA CTCTTGGCCA CTCTTGGTGT CCTCAGTGAA 12780

CAGACCCTGT TGCACTCACT CTCCCTGCCC CACAGCTTTC CTATGGCTCC AGCTCACCAG 12840

CCCTGTCAAA CCGGCAGCGT TTCCCCACTT TCTTCCGAAC GCACCCATCA GCCACACTCC 12900

ACAACCCTAC CCGCGTGAAA CTCTTTGAAA AGTGGGGCTG GAAGAAGATT GCTACCATCC 12960

AGCAGACCAC TGAGGTCTTC ACTTCGGTGA GGAGGGGTTG GGCAAGGGGT AAAGGGACAT 13020

AAGCTCAAAT TCCAGCACCA GGAGATGTGA CGTGAGAGTC ACTTTTAGGG GCAAGAACTT 13080

GATTCTTCAT TGAAAGAGAA CGCATTCCAT GTGGATTAAG TGCAGTTCTT TCTGTAGCCA 13140

GGGGAAAGAA TGAGTTGAGT TTTTGGGATC CTCTCTGTCT TTATGATTTT ATGATTTTTT 13200

TCCCCTGTTT GATGCCCTGT TCCCCAGACA TATAGACCCA GAATGACTCA GTTCTGTTAA 13260

AGTAGGTTCA ATCCAAAGTG GGGGCAAGAG ATGGGAGCGA AGATGAGATA GGAATCCAGG 13320

AAGGCAGCAG ATTCCAGAGG CTTTCAAGGG GGGTGGTGGG TGGGTGTGAA TGGGAACAGA 13380

GGGGATGGAG CCAGTGGATT ACAGAGGAGA GAGGGAGAGG AAAGAGAGAG AGAGAGAGGA 13440

ATGAGGGAGA GGAGAGAGAG GGGCAGAAAG GCAGCTGCAT GGATCTGGTA GTTGGTACTA 13500

AGAGAGAGAA GCCGACAGAC AAGGAGAGGT TGAGGGGGAA GAGGGAGATT TGGGGAGGTA 13560

GAGAGGAAAT ACAGGCTCTA CATCTGAAGA AGGCAGTCTG CTCCCTCCCT TTTATTCTAT 13620

TCTTTGGGTC TTCTATCCAC TGTGTTCAGT GGCCCTTTAA TCCTCCCCCA CTTTCACTCT 13680

GATTCAGACC ATTCTTCTCT GATCCTTTGT CTGTCTGCCC ATTTGCCTCT TGAGGTTGAC 13740

ATCATGCTGT CTGTCCCAGT CCTTGCCTTG TCTTTTCCTG GTTCCTTTAT GTTTCTTTAC 13800

CCCATCTTTG CCTTCAGTGG TAGGAGTGGG TGAATGGAGT GGCTTCCCCC ACACAGAGCC 13860

TCAGCAGGGG CTCACCATTC ACCTTCCCAC TTGGAATCCA CATCCTAAGA CCAGATGCCT 13920

TCCCGAACTC CTCACTTCAG GGACAGAAGC TGTTGAAGGA AGGTTCAGAA TGGCTGCTTC 13980

TTTGCTCTAT CTGAGTATTG CTCTGAAATC CCCAGTTAAC CTCTCTGGTC TTTATTCCCT 14040

CATGCACCCC GTGTTTTTCC AACTTGTTTT TTATTCCCAC CCAAGACTCT GGACGACCTG 14100

GAGGAACGAG TGAAGGAGGC TGGAATTGAG ATTACTTTCC GCCAGAGTTT CTTCTCAGAT 14160

CCAGCTGTGC CCGTCAAAAA CCTGAAGGTC AGATGGCTGG GAGTGGTGGG CTCTGTTTAC 14220

GGAGGGACCA AGCTGGGGGA CAGTGACTGG TTGGAGAGGA AAGCCAGGCG GGGGCAGGTT 14280

TTGATTCTCT GAGGCAATAG CATCTCCTGG GGAAGTTTAG CTCCATCTTC CAGTTGACGT 14340

TTATTCACTA TACGTTGAGC GTTACCCTGC ACTAAGCACT TTGGGATGGG AAATCAAAGC 14400

TGTGAAGACA TCTGGCTTAG CCCCTCAGGC ATTCCCGGGC ATCCCTCAGG AGCTGTTTCT 14460

TTCTCTGTTG TAGCGCCAGG ATGCCCGAAT CATCGTGGGA CTTTTCTATG AGACTGAAGC 14520

CCGGAAAGTT TTTTGTGAGG TGGAGTTGGA TCTGAAGAGG GAGGGGCACT GGGTGGGAGT 14580

TTCCCTTGGT TTTCTTGTGG GGCCTCCTCT TGGCATCTGT GCCTGAGTTG ATAGCATATG 14640

ATCTGAGGTG ACGATTCATA GGATGTCTCT GTCTGTTGGC TCTGACTGCA TCCCTTGTCT 14700

GCACACACAT GATACTTTCT TCAGATCTCA TTTTTCTACT GCTTTGTGTT TCCTGAGAAG 14760

CCCATGAATT CCATCTGTCC TGACTGGCTG GAAAAGGCCA CTCAGAAATA CAGGGGCTGG 14820

GGAGAAACTT AGAAGGAAGA ATTGTCAGCC TTTCCTACTA TCCCCAAGAC TTGTAGATTT 14880

CTCTTTTTAG TTCTACTGCT CTTCCCTGAT TCCCAAGAGG CTAAATAGTA TCAAGTGAGA 14940

TAAGACAAAA ACAAACAAAT GAGCAAACAA AAACTCAGCC ATTCTCCTCT GTATTCAGGT 15000

GTACAAGGAG CGTCTCTTTG GGAAGAAGTA CGTCTGGTTC CTCATTGGGT GGTATGCTGA 15060

CAATTGGTTC AAGATCTACG ACCCTTCTAT CAACTGCACA GTGGATGAGA TGACTGAGGC 15120

GGTGGAGGGC CACATCACAA CTGAGATTGT CATGCTGAAT CCTGCCAATA CCCGCAGCAT 15180

TTCCAACATG GTGAGAGTGT GGGGACTTGC AGTCTGGCAC CTGGGAGGGT GGAGAGGACT 15240

GAGGGGGCCT TGCAGGGGAA AGGGTGGCAG GGAGAGGGTG CGGAATTTGG ATATAAAGGA 15300

GAAGAGGGGG CTGTGCCCAC CCTGAACTTG TCTGCATTAT GTTTCCTGTG GATCCTACCT 15360

TTGCTCTGAC TTCCTTGGGT AGAGAGAGAA AAAAAAAAAA ACGATGGAGT TGTATGTTCA 15420

GTAGGTTCCT GATGAGTGGA AGGGCTGTTA CCATGGAGAC GAGGAGCAGT TGGTGAGAAG 15480

TCAGGAGGAA CCGGCATTAA TGATAATATG GATGCTTGTA TACTCAAGCA CACCTTTACA 15540

GGAGCACTGT GTCTGGGCAG AATTGCATTT CATTTTCTTG GTGATTTATG TTAGTGTTTT 15600

AGAGTTGCTT AATATTCACT CATGATTGAT ATGCAATTAG CTTGGATCCT GTTGCTTTTT 15660

TTAATGATTC TTTTTTTAGA ATTTTATGTG GAGAAGGGGC TTTTGAAATC ATTTAGCCCC 15720

AGACAGCTGG TTAATGACAG ACTTGGAGCG AGGAGGCGCA TTCTCTGATT TGAGAGGGCT 15780

GGGGCAGCCC TGCAGCTGAG GGGCAGAAGT GAGTGTGGGT GTGGGCACAG GGGAAGAGGC 15840

TTCTTTGTTG CTCAGCACTA CATTGTAAAG AGTACAGGAA TTTTGGGTAA GCTCGATACG 15900

GCCACCTATG TGTCTGCTTA TTACACCAAG TGTGTACCTG CTTTGGAGAG CATGTTAGGG 15960

GTGAGTTAAT GGTTGTAAAG CCCCAAAATA GCTTGATTTA AAAAGTTTTA CTAACAGTGG 16020

CATTTGTGTG GTACTTAGAT TCATGATTTC AGGGATTCTA ACAGCAAAAC TCTGATGTGG 16080

ACTTATCTAC ATTTTGTAGA CCAGGAAAAT GAGACTCAGA GTGGTTTAAT GACTTGTCTC 16140

AGGTCATGCA GCTACTCAGC AGTCCAGCTG GAATTGAATC CTAGTTAAAT CTGGCTCCAA 16200

ACTGCTCTGC TCATCACTTC CCATGGAAAA CCATGAAGCC TGACTAAAAA GTTTGTACTT 16260

TTATCTGCTA ACACTGTTTA GCTACTGGCA ATTGCTGAGC AAAGGAGAAA TAAGATTTAG 16320

GAAGTTGACT CTGGTGGTGT GTTGTGATGA AAGATCTGCT GCCACATGGA CTCTGATGAT 16380

ACACACTGTT TTGGGACATT CTAAATGGAC TCTATGTTTT AAGGTATGTT AGTTGAAAAA 16440

TACAATGAAA AAATTTTGGA AGATCACTTT CCTGTAGAGT ACCCAAGAGG TGAGTCATCC 16500

ATAAAACCTG GGAAAGTATT GGAACTGGCT ACCAAAGCCA GAAAGACTTG TTGGTGCCCT 16560

GAGCCCAGGA CTTGCCAATC GTTGCTTCTG TCTTATAGCA GGGAGCTAGA GGAGGCTGAC 16620

AGAAAACAGG GATTTGGCTG GATCTAGTGG CTCACACCTG TAGTCCCCAG CACCTTGGGA 16680

GGCCAAGGCA GAGGATCACT GGAGCCCAGG AGTTCAAGAC CAGCCTGGGC AATATAGTGA 16740

GACCCCCAGC TCTAAACAAA AAGAGAGAGA GAGAGAACAG AGAGATGGTT AGGGACATAG 16800

AGGGACACCT GAGCATTGGG GTGGGAAGAC TTGGAAGGTC CCTTTCTGAC TTTGTGTCTA 16860

GA 16862

(2) INFORMATION FOR SEQ ID NO: 61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16707 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION:1..2742

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:2743..2985

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION:2986..3613

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:3614..3677

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION:3678..5226

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:5227..5304

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION:5305..6848

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:6849..6999

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION:7000..7494

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:7495..7627

(B) LOCATION:11370..11591

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:11592..11735

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION:11736..12512

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:12513..13999

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

AGCTAATTTT GGAATTTTTT TTTTTGAGAC AGAGTTTCAT TCTTATTGCC CAGGCTTGAG 60

TGCAATGGCG CGATCTTGGC TCACAGCAAC CTCCGCCTCC CAGGTTCAAG CCATTCTCCT 120

GCCTCAGCCT CCGGAGTAGC TGGGATTACA GGCATGCGCC ACCACACCCG GCTAATTTTG 180

TATTTTTAGT AGAGACAGGG TTTCTCCATG TTGCTCAGGC TGGTCTCGAA CTCCGGACCT 240

CAGGTGATCC ACCCGCCTCA GCCTCCCAAA GTGCTGGGAT TACAGGCGTG AGCCACCGTG 300

CCCAGCTAAT TTTTGTATTT TTCATAGAGA CAGGGTTTCA CCATGTTGGC CAGGCTGGTC 360

TTGAACTCCT GACCTCATGA TCTGCCCGCC TTGACCTCCC AAAGTCCTGA GATTACAGAT 420

GTGAGCCTCC GTGCCCAGGC CAAGTCTGGC TAATTTTTAA AAAAATTTTG TAGAGTTGGG 480

TCCTCTCTGT TTTGCCCAGT CTTGTCTCAA ACTCCTGGGC TCAAGGAATC CTCCTGCGTT 540

GGCCTCTGAA AATGAAAATG TTGGGATTAC AGGCGTGAGC CCCCTGTGCC TGGCTGCCTT 600

TTTTTTTTTT TTTTTTTAAA GGCAGAGTCT CACTCCATCG CCCAGGCTGA AGTGCAGTGG 660

CGTGATCTCT GCTCACTGCA ACCTCTGCCT CTTGAGTTCA AGCGATTCTT CTACCTCAGC 720

TTCCCAAGTA GCTGGGATTA CGGGAGCCCA CCAACACACC CAGCTAATTT TTGTATTTTT 780

AGTAGAGACG GGTTTCACCA GGTTGGCCAG GCTGTCTGGA ATTCCTGACC TCAGGTGATC 840

CACCCTCCTT GGCCTCCCAA AGAGCTGGGA TTACAAGTGT GAGCCACTGT GCCCAGCCTG 900

ACTTGTTTTT TATAATGCCT TTTTTTTTTT TTTTTGAGAC GGAGTCTTGC TCTGTCGCCC 960

AGGCTGGAGT GTAGTGGCGT CATCTCAGCT CACTGTAACC TCCACCTCCT GGGTTGAAGT 1020

GATTTTCTCA CCTCAGCCCT CAGCCTCCTG AGTAGTTGGG ACTGCAAGTG CACACCACCA 1080

TGCCCAGCTA ATTTTTTGTA TTTTAGTAGA GATGGGGTTT CACCATGTTG CCCAGCTGGT 1140

CTTTAACTCC TGAGCTCAGG CAGTCTGCTT ACCTTGGCCT CCCAAAGTGC TAGGATTAAA 1200

GGTGTGAGCC ACTGTGCCTG GCCTTTTTTT TTTTTTTTTT TTTTTTGAGC AGTTTTAGTT 1260

TCCCAGCAGA ATTGAGATGA AGGTACAGAA ACTTCCCATA TGCTTCCCAC ATGCATAGCC 1320

TTCTACATTA TCGACATCCT CCGCCAGAGT GGTACATTTG TTACAACTGA TGAACCTACA 1380

TTGATACATC ATAATCACCC AAAGTCCATA GTTTACATTA GAGTTCACCC TTGGTGTTAT 1440

ATATTCTATG GGTTTGGACA AATGTATAAT GAGACGTATC TACTATTAAA TACTTTACAG 1500

AGTATTTTCA CTGGCCTAAT CCAATGGACA TTTATTGTTA CTTCATTATG GTTGGGCACA 1560

GTGCTAGATG CTGATGATTA AGAGAGGGCA TGGGATTTGG TCTTGTCCTC AAGGGTAGAA 1620

CCTAGGCCCA TTGCATCTTC AAAGCCCAGG CTCCTTCAAA GCCCAGTGTA GTAGCAACTG 1680

CTGTACCTTG CCTGTGCCCT TTGCGTATCT CACTCCTCTA TCTCTCTAGA AAGTTGGAGA 1740

GAAAAGTGAG CAAGGCATGA GGAACAAAGT TATTTATTTA TTCTTCATTC ATCTATTTAT 1800

TCTTTCATTA CCGTTTGTGT TAAAACATTC CAAACCCAAA CAATTATTTG TATGGTCGCC 1860

TGTGTATTAC TTGTGGTTTC CCAAGAAGTA GTTGCTAAGC TTTTCCTTGT ATGGTTTCTG 1920

TGAGGTAAGG AAGGAATGAT GTGATTTTCT CCAGTATGTA GAATGCAGTT CCAAGAGGTT 1980

AAGTAATTTA CTTACAGTTA TTTAGCCAAA CAAGGTTACT GCAAGGTATA TGAAGTCAGG 2040

TCTCTTGACC CAGTTCATGA GAGAGTTAAA GGAACTATCA TTCTTTTTAG CTTTCATGGA 2100

AAAAGAAGGT TGAGTGTTGG GAGGGGTGTG GGTAGGATTG ATAATGGACT TCAAAAATGT 2160

GAAGGGTATT TCTGTAGTTT TCATTCTTCT GAAAGCCTTC TAAGAGGCAG TGAACCAAAA 2220

GCACACAAGA ATGGCAAGAA GTTAGCATGC TGAAGAAATA TCCTCCTGGC TGGCAAGCAG 2280

AGTGAGAAGA CTGCTATCAC CTTTTCTAGA ATCTTTTGGA ATTGTAGGAG CTGTTAGATC 2340

CTGGGTTAAC TCTATGAAGA AAGTCAGAAG GATCAGAGAA CATCAGTGTC ACAGCTCTTC 2400

ATTGGAATAT CCATGTCTCC TCCTTTACTC TGCTCTACCT TCCATCCTTT GCCACTAATT 2460

ATCCAGAGTG TTTGTCAAAA TTCTCTGTTT GCAGTTCTGA GCTAGCAACT GTACACACTA 2520

ACACCATCAG ACACAGCTAA TACCTACTCT AGTCTAGTAG CTTCCGATCT AAGGCAGACA 2580

CATGGGTATA GTTAAAGATT TTGAATGTAC ATGTGTCCAA TCTGACAACA GTAACACAAA 2640

CCATCCATTC AAGTAGAAGT GATTGAGTCA GAATTGGATT GCACCCCTTC CCCCACACCC 2700

ACACACATTT CAGTTCTTTC CTCATGATTT TTTCCTCCCA AGACATCCCA GGAATTTGTG 2760

GAGAAACTAA CCAAGCGACT GAAAAGACAC CCTGAGGAGA CAGGAGGCTT CCAGGAGGCA 2820

CCGCTGGCCT ATGATGCCAT CTGGGCCTTG GCACTGGCCC TGAACAAGAC ATCTGGAGGA 2880

GGCGGCCGTT CTGGTGTGCG CCTGGAGGAC TTCAACTACA ACAACCAGAC CATTACCGAC 2940

CAAATCTACC GGGCAATGAA CTCTTCGTCC TTTGAGGGTG TCTCTGTGAG TTAAAACTTC 3000

CTTCATACTC CCCTGTCTTC CCAATCTTGA GAGAGACTCC CAAGAGGCAC CTTCTACAAA 3060

CATGCATTCT CTGTTTTTCT CAGTTACTTC TTTGCAGAAT CAGTCTCCGA CCAGAGAAGT 3120

AGGGACCTTC AAATTAGAAG AACCCATCAA AGACTAGAGG AAAAAAAATG ATGTATTCCA 3180

TTTTTTTAAA CCCCTCCCCT CATTTCTTTT CAAACTAGAC CAAGTATTCA TGAGTCAGAT 3240

GAGAACTATA GGATTTTGAA AGACAAAACA GTCTGAAAGG TCATCTTCTT ATTCCTTTTA 3300

AAATGAAAAG ATTAGTTTCC AGAGAGATTT GCTGACTTGC TTAGGCCACA CAACCAGAAG 3360

CCTGCTGGTG TTCTGTCTGG GGATTTTTTC CCATTCAAAT CTCATAAGTG AAGCTCCTTC 3420

TCCAAAGAAT AATGTTTCTA AAATCTAGGG TATGGGCATC TGGGGTATGT CCTATATGCA 3480

GGCAAATGCC ATAAATAGCA TTCATTCAGA GGCTCAATTA CATCAAAAAC AGAAGGATTT 3540

AAAGAGTCCC TGATGTTCTC TTTCACTCTT GCTTTTGTCT CCTTTGCCTT GCTCCACATG 3600

TTCCTTCCCT CAGGGCCATG TGGTGTTTGA TGCCAGCGGC TCTCGGATGG CATGGACGCT 3660

TATCGAGCAG CTTCAGGGTT AGTACAGGGG CAGGAGGGGA CCGGACATGG GGGCTAGGCT 3720

GGGGCTGGGC TGGGATGCCC CCTGGGGAAG AATGCCAGAG ACATCACAAG ATTGCCCTGG 3780

CACCTCCCAA CTTCTGCCCT TCTCTTTTAA CTCTGTTCAT CAAGCTTGTA AATAATAATA 3840

ATAATAAGCT TAACTACAAG AAGATTGATG TCTTTGAGTT GCACTGGTTT TGCTCTTGAA 3900

AAGAGGTGTG CAGGCTGGGT GTGGTGGCTC ACCCCTGTAA TCCCAGCACT TTTGGGAGGC 3960

CAAGGCAGGC AGATCATGAT CATGGTCAGG AGTTTGAGAC CAGCCAGACC AACATGGTGA 4020

AACCTGTCTC TACCAAAAAT ACAAAAAAAA AAAAAAAAAT TAGCTGGGTG TGGTGGCAGG 4080

TGCCTGTAGT CCCAGCTACT TGGGAGGCTG AGGCAGGAGA ATCACTTGAA CCCAGGAGGC 4140

AGAGGTTGCA GTGAGCTGAG ATCACGCCAC TGCACTCCAG CCTGGGTGAT AGAGTAAGAC 4200

TCTGTCTCAA AGAAAAAAGA AAAGAAAAGA GACATGCAAA TTAAAAACAG CTACTCTCTT 4260

TCCCAGTGGC TTCCATTAAT TTCAGGAATT TCCCCTTGAG TGGCTTGGGT TGAGAGGTTG 4320

ATGACCTGTC AGTTAGACTC AAGAAAGCTG AATCTAGGAG AACCGCTATT TTTTTTTTAA 4380

GGGAATCTGC CAAATTTCCT TGCTGTGTAA AGCTTCAATG TGTATAGCTT GGCTTTTGTA 4440

GATTGTATTT TCTTGAAACT TAGCACACAG GTATTTGCAG AACTTCTAGG AGTTAATTTT 4500

TCTGCTCCAC TCGGCTCTCA GTCTTTTACG GCATGGCCAA GAGAGCTATT TCTTGGCCTC 4560

CTGTGAAAAG TTTCTTTCTT CCTTTCTCCC CACCTCCACA TCCTTTCAGC TCCTCTTTGT 4620

ATCCAGGACA AGAGGAAATG GACTTCAGCC ATGGTGAAAG GAGTGTGAGT TGGCTTTTGA 4680

AGGAAAAGTT ATGGTAACGG AAACAGTTCT AGAACAGAAA TCTTAGAAAT GACCAAATTT 4740

TACTCAATGG CGCTTTAAGA GGCAGATATA ACTTATCCAA GGAATTAAAA CCCAAGCCAA 4800

CAGAAGAGAA TGTTCTAAAA TTAAAATGAA AGCCACTGGG AAAATAGAGC CTGCCCATCA 4860

TGAGAGGAAG AATAAGCAGA AATATGTGTA AAGCTTTAGA AGCCAAAATC AAAGTGAGAG 4920

ACATCTCGCC GAGAGAGGTG TGAGGAATGG AATAGGTGGC AGACATGTTG TGGAGCCTCC 4980

TCACTGAAGA CTTTTAAACA TAGATATTCT TATTTATTTG AGTTGTCTTG GGAACCACCT 5040

TATATTGCTT TTAAGTCATG TTGCTGATTC AAGAGTCTCG TAGGTCCTTC CAAGCATCCT 5100

TAGGGCCTCA GGTGAAAATA AAATCAGATA CAACCATGCA AAGCTCTAGG GAAGTGGGAA 5160

GTTGAAAATG CCTAGGATCA GCTCTTTGGC TACCTGTGGT CACTCCTTTT ATTGTCGTCT 5220

GCCCAGGTGG CAGCTACAAG AAGATTGGCT ACTATGACAG CACCAAGGAT GATCTTTCCT 5280

GGTCCAAAAC AGATAAATGG ATTGGTGAGT GGATCTTGTT TGTATTTTCC TTCAGCCCCT 5340

CTCGACAGTC AAGGGAAAAA GTCATGCCTT TGAGTGAGGA TGGAATGGTA GAGACTGTTA 5400

GGTTGGAATG TGGCTGGCAG CTGGGCCAGG AGAAAGGGTT AAGTGAGAGT GAATACAACC 5460

CCTAAGGCGT GGGTAGGGGA GACTGGTGTA TTTGGAGAGG GAATAGGCGG TGGTTAGTAC 5520

TATTTTTAAT GGTGCATTGC TGGGGTAACT GGGGATTAGA GGCAGGGGGT GGGCAGAGGG 5580

CGGGAAATGG AAACTCCATT TGGGTTTCCC AGATGTCCTG GTGTCTTGAT ATATTTGAAC 5640

CAGCTACTTC AAGCCCAGAG CTGTCTCTTT GTCTGTCTCT GTCAGGAAAA CGGTTGCTTA 5700

AACTATGGAG GAGGAGGGAA AACCTCATGT AATTGTCATC TGCCAAAATG TGCTTTTTAT 5760

TTTTATATGT ATTTTTAAAA ATTTTCCTAT TTTTATGTAA TTTAGAGGTA GACGTGCAGT 5820

TGTGTTACAT GAATATATTG CATAGTGGTG AAGTCCGGGC GTTTAGTGTG CCTGTCACCC 5880

GAACAGTGCA CCTTGTACCT AATAGGTAGT ATTACATCCC TCAAAATATA CTTTTTTAAAG 5940

AGAGAAAGCA AGCAGTTATT CTTTGTGTAC TTGGTCTAAA TGATAGGACA TAGGAGAGAA 6000

ACTGAAGGTG GACAAAAGGA AGGACCTACT GATAAAAGAA AGCCTCCTTG AGAATGAAGG 6060

GGAGGCTCAA CCATTGAAGA TGGCTGCCGT CTGCCCTGCC CAGCAGATAT CCAGTCATTC 6120

CCAGCACTGC TGGAGTTTTG CCCTTTTTTT TTTTTTTTTA CAATTCGAAT TTAGGACAAT 6180

GTTCTGGATT GCTATAAATG CTGCATGGCC TAAATTATTC TTTAAAAAAA AAACTAAGCA 6240

AATTGAAATT AGTTTTTTTT GGTGAACTCT GACAAATTTG AACTTCCCCC TAATAATAAC 6300

TGGAAAACAT ATTTGGGAAT ATTACCCTGC CAGGATTAAA ATTTCAGATT AGCTTTCCTT 6360

CTTTTTTTGT TTGTCTTAAG AATAGGTGTC CACACTAGAT ACTTCAAGGC CTTTTTAGCT 6420

TTATGATTCC ATAATTGTCA TTTAAAACTT TGATTTGGGT TATAAGAAAC CTTATAACAT 6480

TTTTTAATGA TCCCCTTCTT TCTCCTCCCA TTTTCCTTTG CTGTAAGAAA GACAGAAAAA 6540

CTTAAAGAAC AAACAAAAAC AAAGACTACA ACTTTGGGGA CATGCCTCAG CATTTCCCAA 6600

CCTATGGATA GACCATTCAC TCCATCTTCT CATCTCATTT CTGGTTGCTT CCTAACGGCC 6660

CCAGTGGCAC TGAGCATTCT GCCTGCAGTA ACCTCTGTCC AGTGCAGTTA GGGCCTCATG 6720

TCCCCAGCCA ATGACTGAAT GTCCATCAGC AATCTAGTTC TTTGCCCTTT TCTCCTATCC 6780

CGTCTTCATT CCTTTGTCCT CCTTCCCTTC TCTTTTCCCT TCCCCTCTTC CTCCCCTGTG 6840

CCATGCAGGA GGGTCCCCCC CAGCTGACCA GACCCTGGTC ATCAAGACAT TCCGCTTCCT 6900

GTCACAGAAA CTCTTTATCT CCGTCTCAGT TCTCTCCAGC CTGGGCATTG TCCTAGCTGT 6960

TGTCTGTCTG TCCTTTAACA TCTACAACTC ACATGTCCGG TAAGTTTCTC TTCTGACGTT 7020

TTCCTTGTCT GCCTCTCTGA GATACTGATC ATGTTCCTGG ACAGGATGAG AATAAAACCT 7080

GTGTAACTCC CATGGCCATG TATCATGGAG TTTTTCATTC TGACTTGTTG AGAATGAAAA 7140

CAGGGAAACC AGATATAACC CCCACTCCTA CTCCAAAGTA GCTAACGGGA GGAAAAAAGA 7200

AAAGAAGAGA AAAAAACAAC CTTTGGGGCC AGGTCTCACA GTCTTGGACT CTACATAAAT 7260

AGCCTGTATT CTAGTGGGGG CCTGTGCTTG GGAAGCCCTC TGCAACTCCA TCTTCAGCCC 7320

CATGACTGCA TTGCTCTGCC TCTCAAGGCT CCACTGTCTT CTCCAATCCT GTCTTCCTTT 7380

AGCCCCTGGC CCTGAAATTA GGGTCATGCC ATTGCGTGGT ATTTGGAGAG CTCAGCCTCC 7440

CTGGAGAAGA GGGGTAATTC TCTCTCCCTC TCACCCTCTC CACCTCTGCC CTAGTTATAT 7500

CCAGAACTCA CAGCCCAACC TGAACAACCT GACTGCTGTG GGCTGCTCAC TGGCTTTAGC 7560

TGCTGTCTTC CCCCTGGGGC TCGATGGTTA CCACATTGGG AGGAACCAGT TTCCTTTCGT 7620

CTGCCAGGTG AGGAGGTGGT GGGCAAATTC CTTACAGGAT GTGACTCTCC CACCCGTCTC 7680

AGGAGCACCT TCCATGATTT ATGATTCTCT GCCCTTCCTC CTCAGCTTTC CCTGACTCTT 7740

GTCCCTGTTC TTTCCTTCTA GCATCACCCC TCTGTTCTCT GTTTGGCTCT GTCCCTTCTT 7800

TCTGTGTCTG CAGGCCATTT TCATTCTGTA GTTTACTTGT CAGTTCCAAG GTTGCCATGG 7860

CAGSCCTYGC AGAGAAGAGG AGGGAGCCAT TGAAGGCAAA GGAAGGGGAT CTGCTCAAAG 7920

GTCTCCTGAA CAATGGTGGC TTGTCTGTGG TATGGGGGCT GAGAATCAGA ACTGTGGACT 7980

TTTTTTGGGA GCCTTTGTTG GGTTTGGAAG GATAGAAGCA GAGATGGAAA CACAGCAGAG 8040

AGTTGGGGGG AAGGGACCAC TGCCACACAG GGGAGGAGGG GCTCTGGGAC TGTTGGTACA 8100

TGGAAGGTTC TAGTGCTGTG GGGAGAGGCC AGCTTCAACA GTGATAGTTG AGTGGTTCTC 8160

TTTTCCACTG GTGGAAACAC CCACTCTTTC TCCTGATCTG CCTGCCTGTC CTTGCTCTCT 8220

CTTTTTCCTC TGCTCTGTGC TGTCCTGATC ATACATCTGT GCACATGGCA TTTCCATGCA 8280

CATGCACATG CAGTTCATCA GGAATCCTCT GTTCCCAGTG AGGCCAGAGT GCAGCTGGAG 8340

AAGCAGACAA TTAGCTGTAG TGCAATAGGA GAGGTTCCAG AGTAGGGATC TGCACAAAGT 8400

GCTTTGGGGG CAAAGAAGGG AACACAGTTC ACTGCTGGCG TGATTGGGTG GACCTCACTG 8460

AAGAGGTGGC ATTTGAATAC TGAAGGACAA ATAGGATTTT ATCAGCTAGA GAAATAGAGG 8520

AAGGCTACTT CAGGGGCATA GGGAGCATCG TGTGGCTAGA AAATACATGA AAGAGAGTAG 8580

ATGAAGAGAA AGTGAGTAGT TCAGCATGGC TGGAGCGTGG GGTAGGTGTG GGGCTGGGAG 8640

ATGAGCCTAG CTGGACAGGT GGATGGGAGC ATGTTGTGAA GGGTCTGTGT CATATCCAGA 8700

AGTGTTCAGG CTATAACTTA TAGATATTGG GGAGTGGTTG GAGGTTTTTG GCCACTAAAG 8760

CCAGGAGGTT TTAGCAAGAT CACCCTGGTG GTGTGGAAGT AGAGGGTGGA TGGGAGGAAT 8820

TGTTCAAGGT GGGGAGACTG CTCTCCTCCT GCCGCTCCCC GTCCTGCTCA CATTTTCGCA 8880

TCCTCCCTGT GCCACCATGA GCTCCCTGCC CGTGCTCCCT GCCCACTCTC CCTTAGGGTT 8940

CTGCCCATCC TTACTGCAGT CCCGGCTACT ACTCTACCCT GTTCTGCCTG TGCCCTCTCT 9000

TCCTTTCTAG GCCCGCCTCT GGCTCCTGGG CCTGGGCTTT AGTCTGGGCT ACGGTTCCAT 9060

GTTCACCAAG ATTTGGTGGG TCCACACGGT CTTCACAAAG AAGGAAGAAA AGAAGGAGTG 9120

GAGGAAGGTG AGCTGCTGCC CAATCCTCAG CCCCCARATC CTTGGCTCCT GGGGCACAGA 9180

GCATTTTCCC CTGACGTGCC TGTTCTCCCC ACATATTTAT CCAGACTCTG GAACCCTGGA 9240

AGCTGTATGC CACAGTGGGC CTGCTGGTGG GCATGGATGT CCTCACTCTC GCCATCTGGC 9300

AGATCGTGGA CCCTCTGCAC CGGACCATTG AGGTACCACT GGAGAGGAGG TGCTATGGTC 9360

AGGAGAATGA GCAGGGCTCA GTGGCCATCA GGGCCCTGGG GCTGTGTGTG TCTTGAGGGA 9420

TGAAGCTACT TGGAGAGAGT GCCTTCCTCG TATTGGAAGC TCTTCCTTTC CTTCCTAGAA 9480

GGAGCCCCTC ATAGGCCTCC AGATTCAGCT GAAGAAAGGA AGGGGTGGGA ATCTGGGAAG 9540

GGTGTGTAGA ACTTCCAGGC ATCAGGGAAA GTGGGGAACA AGCACCTCCA AGGGTTCAGG 9600

AAAACATTCT TAGGCCTAGA ATGAGATTTG GCATCAGCAT TGAGGGTCTC ATAGGAAAAC 9660

AGTTGGAAGC CAGAGACTGA GAAGCGTTGA GGAGAGGAGG GGAGGCTGGC AACCATCTTT 9720

CTTGTGACCT TGTTTCTGCC CTAGACATTT GCCAAGGAGG AACCTAAGGA AGATATTGAC 9780

GTCTCTATTC TGCCCCAGCT GGAGCATTGC AGCTCCAGGA AGATGAATAC ATGGCTTGGT 9840

GTGTGGGATG TGGGCAAAGG AGGGCAGGGA TGCACAAAGG CAGGAGGGAA GGCAGGGGTA 9900

GAGGGCTTGG AGGGAGAGGG GTCTTTGGAA GAGGAGGTAG AGAGCTTGTC AACCCAGTTT 9960

GAACACCCTA CTCTTTGTTA TTGCACTAAT CTTTTCTGAG AATAGGGGAG AGTTGCTCTT 10020

TTGCTATGAG GAGCTTAGGG CCCAAAGCAC AGAAAGCACA GATGAAGAAC TTGTGTTCAG 10080

CAGAGGAACA AGTGGGGGTA ACCCCACCTC CAGACTTGAC ATTATCTTTT AGATCCCCCT 10140

TGGCCTTATT AGCATTGTTC GATTCATGGT CACAAATTGC AAACCTACCC TCTGCCTGGA 10200

AAGCCACCTT CCCACCTGTA GGGTAAGGGT GAGGCATGTG TGGCCCAGAC TGGCCTATTT 10260

CTAGATATTC AACAAGCCCT TGCCTGACTG ACAGCAGCTT GCCACCATTG CTTTCCTGTG 10320

TGAATCCCAG GAAAAAGTGA TGTGGTCTGG GCAAGTTGGG TGGACATAAG GGATAGGGGA 10380

CACAGGGTGA GGTTTGCTAG GTCAGAGGGG TTGGATTGGA GAGGAGGGCC CCCTTTCCAT 10440

TTCAGAGTAG GTGAAGGGCA GAGAGGGGAT GGGGATTGAG TGAGGAGCAT TGTGGTCCTT 10500

GTTGCTCAAG TGACTCTCTC CTGCCATCCT AGGCATTTTC TATGGTTAAC AAGGGGCTGC 10560

TGCTGCTGCT GGGAATCTTC CTTGCTTATG AGACCAAGAG TGTGTCCACT GAGAAGATCA 10620

ATGATCACCG GGCTGTGGGC ATGGCTATCT ACAATGTGGC AGTGAGCACT GACCCCATGG 10680

CATTGACCCT GTAGGCTGAC CACAGCAGCC CAGATATAGA GGACTAGGAA GAATCAATGC 10740

TAGATCTGGG ATCGGTTGCT TAGAAGTCTT AAAAAGTTTG TTAATTCTTC AGGTCTATAA 10800

AGCACTTTAC AGTTTACAAA GCTCACTACA GACATTGTAT CATTAATCTT GCAACTACCC 10860

AGTGAAGTAG ATATTAGTAT CCCCACTTTA TAGGTGAGGA AACAGAAACA CAGAGACGTT 10920

AAATTGCTTG TCTGTGGTTA ATGGGCTGGA CTCTATTGAC ATTTCCTGCC AGGGACCGAC 10980

TCTGGAGGAC CCGGAATCTG TGCATAGAGA TCCTGGGAGT TCCTGCCTTG AGGGGAGGGG 11040

TTAACCAAGA GTGAAAACTG GTTTGGGACA GTTTGAGATT TTTCTCAATC TATATTGAGG 11100

ATGATCCTGA ATTTGGATCC TTTTCAAAGG GAAAGTTCAC CAGGAAACTG TCTGCATAGA 11160

CTCCCTCCCA TGGGAAGTAA ACTCTGGATC TTGTCTGAGC CTGCAGACCT GAGACTCCCT 11220

CAATGTGTCT TTCCCTCTAG GTCCTGTGCC TCATCACTGC TCCTGTCACC ATGATTCTGT 11280

CCAGCCAGCA GGATGCAGCC TTTGCCTTTG CCTCTCTTGC CATAGTTTTC TCCTCCTATA 11340

TCACTCTTGT TGTGCTCTTT GTGCCCAAGG TAAGGATCTG GCTTTTCTCC CACCCTCTTT 11400

GTTCCCATGT TCCCTCCATC CCTCCTTCCT ATATTACTGA GTTCCTCTGC CCTTCCGTTC 11460

ACCCTCCTCT CACTCCTCCC CTTGTTTTGG GCCCAACTCT TATCAGCATT CCTTCCACCT 11520

CCAACCTTCC ATCAGCCAGT CACTAGTACA GTCCTTGCTG GGCCACCCCA CGCCCAAACA 11580

TTTGCCCCCA GATGCGCAGG CTGATCACCC GAGGGGAATG GCAGTCGGAG GCGCAGGACA 11640

CCATGAAGAC AGGGTCATCG ACCAACAACA ACGAGGAGGA GAAGTCCCGG CTGTTGGAGA 11700

AGGAGAACCG TGAACTGGAA AAGATCATTG CTGAGGTGCG GGGGTGGGTG TCAGGGTAGG 11760

GTGTTGGAGT GGTCCAGGAG GCTTGCGTCT TAGCTTGGGT TGTCTGAAGC CAAGCCTGAG 11820

ATACAGGGTC AGATGTTCTT GGCTCATGGA GGGAGGGTCC TAGGAGACAA CCTGTAAGGA 11880

GTGAATGGAG CAGCATAGGG GAGGGGAAAG GGCTGAGCAA GATTCTATCT CAGGCAAAAT 11940

CCAGTGTTGG CCTGGCAGGT GGAAGGGCTC TGGAGTGGGA GCTATGTGGT TGACTCAGCC 12000

TCCTTAAGGC AAGAGGATGG CTGTTGGCTG TAGGTGACAA CTGGAGAGAG GCAGCTGTGA 12060

GCCTCTAGTA GTCAACACTC ACAGCAGCTG GGTGTAGCAT GCAGCCCCAG CATAAAGGAC 12120

CTGGGCAGGC GTTCACTGTG CCCCAGGCTG TCATTAGGGG CTGGTGCAAT GCCAAAGAGA 12180

GGGATGTTCC AACTGGGTTG ACACATCTCT CTGATTTATT GGAAGCTCTG TGCACTGACT 12240

TTTCTCTCCT TCCCCACTTT TTCCTTTTGT TTTTAAATTC TCTCTTATTT CCCTGATCGC 12300

ATTTTTTCTA TCGGTATCCT TATGTTCTCT GGCTTTTCTT GTTCTGTTTT GATTTCTCCT 12360

TTTAATTTAT TCTGTCCACT TACCCTACGT CCTCCCCCTA CATTTTTCTG TGCCCTTCCT 12420

CTCTTTCCCT GTGCCCTTCC TCTCTTTCCC TCCTCCCCAC TCCTTCATCA CCTCCTCTTC 12480

TCCTACTATC CCAATTGTGC TTCTTCCTCC AGAAAGAGGA GCGTGTCTCT GAACTGCGCC 12540

ATCAACTCCA GTCTCGGCAG CAGCTCCGCT CCCGGCGCCA CCCACCGACA CCCCCAGAAC 12600

CCTCTGGGGG CCTGCCCAGG GGACCCCCTG AGCCCCCCGA CCGGCTTAGC TGTGATGGGA 12660

GTCGAGTGCA TTTGCTTTAT AAGTGAGGGT AGGGTGAGGG AGGACAGGCC AGTAGGGGGA 12720

GGGAAAGGGA GAGGGGAAGG GCAGGGGACT CAGGAAGCAG GGGGTCCCCA TCCCCAGCTG 12780

GGAAGAACAT GCTATCCAAT CTCATCTCTT GTAAATACAT GTCCCCCTGT GAGTTCTGGG 12840

CTGATTTGGG TCTCTCATAC CTCTGGGAAA CAGACCTTTT TCTCTCTTAC TGCTTCATGT 12900

AATTTTGTAT CACCTCTTCA CAATTTAGTT CGTACCTGGC TTGAAGCTGC TCACTGCTCA 12960

CACGCTGCCT CCTCAGCAGC CTCACTGCAT CTTTCTCTTC CCATGCAACA CCCTCTTCTA 13020

GTTACCACGG CAACCCCTGC AGCTCCTCTG CCTTTGTGCT CTGTTCCTGT CCAGCAGGGG 13080

TCTCCCAACA AGTGCTCTTT CCACCCCAAA GGGGCCTCTC CTTTTCTCCA CTGTCATAAT 13140

CTCTTTCCAT CTTACTTGCC CTTCTATACT TTCTCACATG TGGCTCCCCC TGAATTTTGC 13200

TTCCTTTGGG AGCTCATTCT TTTCGCCAAG GCTCACATGC TCCTTGCCTC TGCTCTGTGC 13260

ACTCACGCTC AGCACACATG CATCCTCCCC TCTCCTGCGT GTGCCCACTG AACATGCTCA 13320

TGTGTACACA CGCTTTTCCC GTATGCTTTC TTCATGTTCA GTCACATGTG CTCTCGGGTG 13380

CCCTGCATTC ACAGCTACGT GTGCCCCTCT CATGGTCATG GGTCTGCCCT TGAGCGTGTT 13440

TGGGTAGGCA TGTGCAATTT GTCTAGCATG CTGAGTCATG TCTTTCCTAT TTGCACACGT 13500

CCATGTTTAT CCATGTACTT TCCCTGTGTA CCCTCCATGT ACCTTGTGTA CTTTCTTCCC 13560

TTAAATCATG GTATTCTTCT GACAGAGCCA TATGTACCCT ACCCTGCACA TTGTTATGCA 13620

CTTTTCCCCA ATTCATGTTT GGTGGGGCCA TCCACACCCT CTCCTTGTCA CAGAATCTCC 13680

ATTTCTGCTC AGATTCCCCC CATCTCCATT GCATTCATGT ACTACCCTCA GTCTACACTC 13740

ACAATCATCT TCTCCCAAGA CTGCTCCCTT TTGTTTTGTG TTTTTTTGAG GGGAATTAAG 13800

GAAAAATAAG TGGGGGCAGG TTTGGAGAGC TGCTTCCAGT GGATAGTTGA TGAGAATCCT 13860

GACCAAAGGA AGGCACCCTT GACTGTTGGG ATAGACAGAT GGACCTATGG GGTGGGAGGT 13920

GGTGTCCCTT TCACACTGTG GTGTCTCTTG GGGAAGGATC TCCCCGAATC TCAATAAACC 13980

AGTGAACAGT GTGACTCGGC ACCTTGCAGT CTTCCTGTGA ACAGAATGGG CTTCAATCCA 14040

AGAAGGGAGG CTCAGAGGAC TCCAAGTTCA TGAAAAGGCA TTAAAGCGGA GGGTGAAAAG 14100

AGGTGTTTTA TTGATCCATT GAGGGCTTAG CAGAATGAAG CAGGACATGA TTAAGTCTGA 14160

GATTAGTGAG TGAGGACACT ACTGGTTAAA AGTGTGGGCT CTGGAGTCAG ACTGCCAGGG 14220

TATCAGATCC AACCACATGC AAACATTTTC TTAGTCTCTA TTCCCCATGT CCTCATTTAT 14280

GAAAATGAGA ATAACAGTAA TACATTCCTC CATAGGTTGG GTACAAAGAC TATTATAAAT 14340

TGTGCATTCA GGTGCCTAGG TTGGCCCTTG GGCCATGGTA TATGTTGCGT GAATGTTAGC 14400

CTCTGTCCCT GCTGTTTAAT GAGTTCCTTG ACAGTAGTGG GCATGTATTG GGAGCCTGGA 14460

GCAAGTGCCT AAGCATCCCC TCTAGGGACG CTCCTTCCCA GGAACTAAGA AGAGTAAAAG 14520

AATGATGACT GCTAGAAGGT AATGGATGAG ATGGCTGCTG AGTGCTTCCA ACCTTAAACA 14580

TCTTTGTTTA GGAACTCTGA GCATCTTGGA AATAATTTGC TATCAAACTG AAAAAAAATC 14640

TTGAATGGAC AAGGGCAAAA ACATTTGCCT GAGACTTTAA ACATTTTTTG TGTCATCTTG 14700

GAGAGTGCTT TTTTGAAGCT CAAATTTTCT TTTATTTTGG CACTGATTTT TAAAGTGATT 14760

CTCAGATTTT TGTAAGACAG CTGCAAGGGT TGGATGGGCC CTGTCATTCA CTGACCTGTT 14820

AAGAGCCAAT TTCTGAACTT CCACTAAAGC ATGGGCTGGT TGAATCTTAG ACCAGTACTT 14880

GAAAAACTTT CCACTGTGAT TATCCACCTT GGACCAGTTG GACTTAATTT GAGTCTCTCT 14940

TCTTCCCACA GTGAAATATC CGCTAGGAAA AGAGAAGAAG TCTATGGAGT AGATAAGCCT 15000

GCAGTTTGGA AAATGAATAG TTGGCATCAG AACAGCAGCA GGAGGCTGGG TGTGGTGGCT 15060

CACGCCTGTA ATCCCAGCAC TTTGGGAGGC TGAGGTGGGT GGATCACGAG GTCAGGAGTT 15120

CAAGACCAGC CTGGCCAAGA TGGCGAAACC CTGTCTCTAC TAAAAATACA AAAATTAGCT 15180

GGGCGTGATG GTGGGTGCCT GTAATCTCAG CTACTCGGGA GGCTGAGGCA GATAATTGCT 15240

TGAACCCAGG AGATGAAGGT TGCAGTGAGT TTGTGCCACT GCACTCCAGC CTGGGCGACA 15300

GAGTGAGACT GTCTCAAAAA AAAAAAAAAA GAAAAAAAAA AAAGAAACAA CAGTAGCAGG 15360

AGCTATAGAA CAGCCCTGGG TAGAACCTAA AAGACCCAAA TTATCATCTC AAACTTGCAT 15420

TGCACTTAAG TGGGCTGTAA ATTATAAACA AAGGGTGAAA AGTTCTACTG TGGCCAAAGG 15480

TAAGCCAGAC ACTCTGCTAG CAGGAGTGCA GGAGTCGAGA GCCAAACGGT GCGGCTAGCA 15540

GAGTGCCCAG TGCAAAGGGG TGGGAAGGAG TGAGATTGAG AATATTAAAA AGGTACTTAG 15600

AAGAGAACTT GTAAGATTTT TACTGGCCAA ATTTAAAACA TGACTGAGCA CTATTTTTCA 15660

TACAGGCCTC CTACTAATAA GAAAACAATT TTGAGATAAC TACTTATTTG AGTTCACAGT 15720

TAATGTTCCT GATGATTAAG ATCAGTTGCA AATGTTCATC TGTCAATGCT TATCTACAAT 15780

GAGACTTCAT GTATTCATTT CTGAAAGTGT CTTTTCAGGG TGAATGGTGC TATTGATTAG 15840

CACTAATACT AATTATTAGT ACATTATATA TAATTACAAT GAGATATACA CACACATACA 15900

CACACACATA TATACATATA CACACACACA ATGATATATA ATTAATATAT AATTGTACCC 15960

CAAGGGGTGC AAAGGAGATG TGTTGCCAGG TGGAGAGGCT ACCGCTTGGC AGTTCTGGGA 16020

GGACTTGCTC CCTGTGCACT GTGAGGCAGG CTTTGCCTTT CAGACCTGCC TTTGGGTAGG 16080

GTTCAGATCA CTTTCTAACT CTGGAATGTC CTAGAATGTA GACTGCCTGC AGGCTTCCAT 16140

GTCCCCTGCT TTTCCCTGAC TTAGCCTGTT GCCTCCCTGC CCTCCTGTTG GTTGTCTACC 16200

AGTAGAGAGC ACTTTGTGTG CACTTGGCTG CTACATTAGT TAGGTGATCT TCAACAAGTG 16260

TTGATGGTGA GTTGCTGTGG CAGGTGCTTT TTTTGGCACT GAGGCAAAAT GGTGAGTAAG 16320

ATGGCTTTCA AGCGTTGTAC CTTCTCGACG TGGGAAGATG ACCAGTAAGC AGAAAAACAA 16380

ACGAGATCAC TTAAGAGAGC AACCAGGAGT GTTGGGCATC TCACAGCCAT TAGCTCTGGT 16440

GTGAAGGACA AATCTAAAAG CAAGGGGACT GTGTGTTCAT TTTCTGGGGT CACAAAACTA 16500

AGGAGCAAAG CCAGTATTCA AACTGCATAT ATTATTATCT ATTGCCACAA AACCTGTTAC 16560

CCTAAATGGC TTCAGACAAC AATAATCATT TATTATCCAT CATGGTTTTT GTGGGCTAGA 16620

AATTCAGATA GGGTACAGTG GGGAGGGTTC ACCCTGCTCC ACAGCATCTG GAAGATCTGA 16680

AGTCTGAGGG TTGTTGTCCA AAAGATC 16707

(2) INFORMATION FOR SEQ ID NO: 62:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:

ATGCGCGCCG GCAGCCAACA TGCTGCTGCT GCTGCTGGTG CCTCTCTTCC 50

(2) INFORMATION FOR SEQ ID NO: 63:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:

GGTCATCCAG CGTTGAGGTG AAGAC 25

(2) INFORMATION FOR SEQ ID NO: 64:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 64:

GAAGGTTGCC AGATTATACA TCCGC 25

(2) INFORMATION FOR SEQ ID NO: 65:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 65:

CCACGATGAT TCGAGCATCT TGACG 25

(2) INFORMATION FOR SEQ ID NO: 66:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 66:

CTGGTTCCTC CCAATGTG 18

(2) INFORMATION FOR SEQ ID NO: 67:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 67:

CCAGTGGACT ATGAGATTGA G 21

(2) INFORMATION FOR SEQ ID NO: 68:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 68:

CTGGTTCCTC CCAATGTG 18

(2) INFORMATION FOR SEQ ID NO: 69:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 69:

CCAGTGGACT ATGAGATTGA G 21

(2) INFORMATION FOR SEQ ID NO: 70:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2700 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..2700

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70:

```
ATG TTG CTG CTG CTG CTA CTG GCG CCA CTC TTC CTC CGC CCC CCG GGC        48
Met Leu Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
 1               5                  10                  15

GCG GGC GGG GCG CAG ACC CCC AAC GCC ACC TCA GAA GGT TGC CAG ATC        96
Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
            20                  25                  30

ATA CAC CCG CCC TGG GAA GGG GGC ATC AGG TAC CGG GGC CTG ACT CGG       144
Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
```

35 40 45

GAC CAG GTG AAG GCT ATC AAC TTC CTG CCA GTG GAC TAT GAG ATT GAG 192
Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu
50 55 60

TAT GTG TGC CGG GGG GAG CGC GAG GTG GTG GGG CCC AAG GTC CGC AAG 240
Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys
65 70 75 80

TGC CTG GCC AAC GGC TCC TGG ACA GAT ATG GAC ACA CCC AGC CGC TGT 288
Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys
85 90 95

GTG AAT CGA ACG CCA CAC TCA GAA CGG CGC GCA GTG TAC ATC GGG GCA 336
Val Asn Arg Thr Pro His Ser Glu Arg Arg Ala Val Tyr Ile Gly Ala
100 105 110

CTG TTT CCC ATG AGC GGG GGC TGG CCA GGG GGC CAG GCC TGC CAG CCC 384
Leu Phe Pro Met Ser Gly Gly Trp Pro Gly Gly Gln Ala Cys Gln Pro
115 120 125

GCG GTG GAG ATG GCG CTG GAG GAC GTG AAT AGC CGC AGG GAC ATC CTG 432
Ala Val Glu Met Ala Leu Glu Asp Val Asn Ser Arg Arg Asp Ile Leu
130 135 140

CCG GAC TAT GAG CTC AAG CTC ATC CAC CAC GAC AGC AAG TGT GAT CCA 480
Pro Asp Tyr Glu Leu Lys Leu Ile His His Asp Ser Lys Cys Asp Pro
145 150 155 160

GGC CAA GCC ACC AAG TAC CTA TAT GAG CTG CTC TAC AAC GAC CCT ATC 528
Gly Gln Ala Thr Lys Tyr Leu Tyr Glu Leu Leu Tyr Asn Asp Pro Ile
165 170 175

AAG ATC ATC CTT ATG CCT GGC TGC AGC TCT GTC TCC ACG CTG GTG GCT 576
Lys Ile Ile Leu Met Pro Gly Cys Ser Ser Val Ser Thr Leu Val Ala
180 185 190

GAG GCT GCT AGG ATG TGG AAC CTC ATT GTG CTT TCC TAT GGC TCC AGC 624
Glu Ala Ala Arg Met Trp Asn Leu Ile Val Leu Ser Tyr Gly Ser Ser
195 200 205

TCA CCA GCC CTG TCA AAC CGG CAG CGT TTC CCC ACT TTC TTC CGA ACG 672

```
Ser Pro Ala Leu Ser Asn Arg Gln Arg Phe Pro Thr Phe Phe Arg Thr
    210                 215                 220

CAC CCA TCA GCC ACA CTC CAC AAC CCT ACC CGC GTG AAA CTC TTT GAA      720
His Pro Ser Ala Thr Leu His Asn Pro Thr Arg Val Lys Leu Phe Glu
225                 230                 235                 240

AAG TGG GGC TGG AAG AAG ATT GCT ACC ATC CAG CAG ACC ACT GAG GTC      768
Lys Trp Gly Trp Lys Lys Ile Ala Thr Ile Gln Gln Thr Thr Glu Val
                245                 250                 255

TTC ACT TCG ACT CTG GAC GAC CTG GAG GAA CGA GTG AAG GAG GCT GGA      816
Phe Thr Ser Thr Leu Asp Asp Leu Glu Glu Arg Val Lys Glu Ala Gly
            260                 265                 270

ATT GAG ATT ACT TTC CGC CAG AGT TTC TTC TCA GAT CCA GCT GTG CCC      864
Ile Glu Ile Thr Phe Arg Gln Ser Phe Phe Ser Asp Pro Ala Val Pro
        275                 280                 285

GTC AAA AAC CTG AAG CGC CAG GAT GCC CGA ATC ATC GTG GGA CTT TTC      912
Val Lys Asn Leu Lys Arg Gln Asp Ala Arg Ile Ile Val Gly Leu Phe
    290                 295                 300

TAT GAG ACT GAA GCC CGG AAA GTT TTT TGT GAG GTG TAC AAG GAG CGT      960
Tyr Glu Thr Glu Ala Arg Lys Val Phe Cys Glu Val Tyr Lys Glu Arg
305                 310                 315                 320

CTC TTT GGG AAG AAG TAC GTC TGG TTC CTC ATT GGG TGG TAT GCT GAC     1008
Leu Phe Gly Lys Lys Tyr Val Trp Phe Leu Ile Gly Trp Tyr Ala Asp
                325                 330                 335

AAT TGG TTC AAG ATC TAC GAC CCT TCT ATC AAC TGC ACA GTG GAT GAG     1056
Asn Trp Phe Lys Ile Tyr Asp Pro Ser Ile Asn Cys Thr Val Asp Glu
            340                 345                 350

ATG ACT GAG GCG GTG GAG GGC CAC ATC ACA ACT GAG ATT GTC ATG CTG     1104
Met Thr Glu Ala Val Glu Gly His Ile Thr Thr Glu Ile Val Met Leu
        355                 360                 365

AAT CCT GCC AAT ACC CGC AGC ATT TCC AAC ATG ACA TCC CAG GAA TTT     1152
Asn Pro Ala Asn Thr Arg Ser Ile Ser Asn Met Thr Ser Gln Glu Phe
    370                 375                 380
```

```
GTG GAG AAA CTA ACC AAG CGA CTG AAA AGA CAC CCT GAG GAG ACA GGA      1200
Val Glu Lys Leu Thr Lys Arg Leu Lys Arg His Pro Glu Glu Thr Gly
385                 390                 395                 400

GGC TTC CAG GAG GCA CCG CTG GCC TAT GAT GCC ATC TGG GCC TTG GCA      1248
Gly Phe Gln Glu Ala Pro Leu Ala Tyr Asp Ala Ile Trp Ala Leu Ala
                405                 410                 415

CTG GCC CTG AAC AAG ACA TCT GGA GGA GGC GGC CGT TCT GGT GTG CGC      1296
Leu Ala Leu Asn Lys Thr Ser Gly Gly Gly Gly Arg Ser Gly Val Arg
            420                 425                 430

CTG GAG GAC TTC AAC TAC AAC AAC CAG ACC ATT ACC GAC CAA ATC TAC      1344
Leu Glu Asp Phe Asn Tyr Asn Asn Gln Thr Ile Thr Asp Gln Ile Tyr
        435                 440                 445

CGG GCA ATG AAC TCT TCG TCC TTT GAG GGT GTC TCT GGC CAT GTG GTG      1392
Arg Ala Met Asn Ser Ser Ser Phe Glu Gly Val Ser Gly His Val Val
    450                 455                 460

TTT GAT GCC AGC GGC TCT CGG ATG GCA TGG ACG CTT ATC GAG CAG CTT      1440
Phe Asp Ala Ser Gly Ser Arg Met Ala Trp Thr Leu Ile Glu Gln Leu
465                 470                 475                 480

CAG GGT GGC AGC TAC AAG AAG ATT GGC TAC TAT GAC AGC ACC AAG GAT      1488
Gln Gly Gly Ser Tyr Lys Lys Ile Gly Tyr Tyr Asp Ser Thr Lys Asp
                485                 490                 495

GAT CTT TCC TGG TCC AAA ACA GAT AAA TGG ATT GGA GGG TCC CCC CCA      1536
Asp Leu Ser Trp Ser Lys Thr Asp Lys Trp Ile Gly Gly Ser Pro Pro
            500                 505                 510

GCT GAC CAG ACC CTG GTC ATC AAG ACA TTC CGC TTC CTG TCA CAG AAA      1584
Ala Asp Gln Thr Leu Val Ile Lys Thr Phe Arg Phe Leu Ser Gln Lys
        515                 520                 525

CTC TTT ATC TCC GTC TCA GTT CTC TCC AGC CTG GGC ATT GTC CTA GCT      1632
Leu Phe Ile Ser Val Ser Val Leu Ser Ser Leu Gly Ile Val Leu Ala
    530                 535                 540

GTT GTC TGT CTG TCC TTT AAC ATC TAC AAC TCA CAT GTC CGT TAT ATC      1680
Val Val Cys Leu Ser Phe Asn Ile Tyr Asn Ser His Val Arg Tyr Ile
545                 550                 555                 560
```

```
CAG AAC TCA CAG CCC AAC CTG AAC AAC CTG ACT GCT GTG GGC TGC TCA      1728
Gln Asn Ser Gln Pro Asn Leu Asn Asn Leu Thr Ala Val Gly Cys Ser
                565                 570                 575

CTG GCT TTA GCT GCT GTC TTC CCC CTG GGG CTC GAT GGT TAC CAC ATT      1776
Leu Ala Leu Ala Ala Val Phe Pro Leu Gly Leu Asp Gly Tyr His Ile
            580                 585                 590

GGG AGG AAC CAG TTT CCT TTC GTC TGC CAG GCC CGC CTC TGG CTC CTG      1824
Gly Arg Asn Gln Phe Pro Phe Val Cys Gln Ala Arg Leu Trp Leu Leu
        595                 600                 605

GGC CTG GGC TTT AGT CTG GGC TAC GGT TCC ATG TTC ACC AAG ATT TGG      1872
Gly Leu Gly Phe Ser Leu Gly Tyr Gly Ser Met Phe Thr Lys Ile Trp
    610                 615                 620

TGG GTC CAC ACG GTC TTC ACA AAG AAG GAA GAA AAG AAG GAG TGG AGG      1920
Trp Val His Thr Val Phe Thr Lys Lys Glu Glu Lys Lys Glu Trp Arg
625                 630                 635                 640

AAG ACT CTG GAA CCC TGG AAG CTG TAT GCC ACA GTG GGC CTG CTG GTG      1968
Lys Thr Leu Glu Pro Trp Lys Leu Tyr Ala Thr Val Gly Leu Leu Val
                645                 650                 655

GGC ATG GAT GTC CTC ACT CTC GCC ATC TGG CAG ATC GTG GAC CCT CTG      2016
Gly Met Asp Val Leu Thr Leu Ala Ile Trp Gln Ile Val Asp Pro Leu
            660                 665                 670

CAC CGG ACC ATT GAG ACA TTT GCC AAG GAG GAA CCT AAG GAA GAT ATT      2064
His Arg Thr Ile Glu Thr Phe Ala Lys Glu Glu Pro Lys Glu Asp Ile
        675                 680                 685

GAC GTC TCT ATT CTG CCC CAG CTG GAG CAT TGC AGC TCC AGG AAG ATG      2112
Asp Val Ser Ile Leu Pro Gln Leu Glu His Cys Ser Ser Arg Lys Met
    690                 695                 700

AAT ACA TGG CTT GGC ATT TTC TAT GGT TAC AAG GGG CTG CTG CTG CTG      2160
Asn Thr Trp Leu Gly Ile Phe Tyr Gly Tyr Lys Gly Leu Leu Leu Leu
705                 710                 715                 720

CTG GGA ATC TTC CTT GCT TAT GAG ACC AAG AGT GTG TCC ACT GAG AAG      2208
Leu Gly Ile Phe Leu Ala Tyr Glu Thr Lys Ser Val Ser Thr Glu Lys
```

725 730 735

ATC AAT GAT CAC CGG GCT GTG GGC ATG GCT ATC TAC AAT GTG GCA GTC 2256
Ile Asn Asp His Arg Ala Val Gly Met Ala Ile Tyr Asn Val Ala Val
740 745 750

CTG TGC CTC ATC ACT GCT CCT GTC ACC ATG ATT CTG TCC AGC CAG CAG 2304
Leu Cys Leu Ile Thr Ala Pro Val Thr Met Ile Leu Ser Ser Gln Gln
755 760 765

GAT GCA GCC TTT GCC TTT GCC TCT CTT GCC ATA GTT TTC TCC TCC TAT 2352
Asp Ala Ala Phe Ala Phe Ala Ser Leu Ala Ile Val Phe Ser Ser Tyr
770 775 780

ATC ACT CTT GTT GTG CTC TTT GTG CCC AAG ATG CGC AGG CTG ATC ACC 2400
Ile Thr Leu Val Val Leu Phe Val Pro Lys Met Arg Arg Leu Ile Thr
785 790 795 800

CGA GGG GAA TGG CAG TCG GAG GCG CAG GAC ACC ATG AAG ACA GGG TCA 2448
Arg Gly Glu Trp Gln Ser Glu Ala Gln Asp Thr Met Lys Thr Gly Ser
805 810 815

TCG ACC AAC AAC AAC GAG GAG GAG AAG TCC CGG CTG TTG GAG AAG GAG 2496
Ser Thr Asn Asn Asn Glu Glu Glu Lys Ser Arg Leu Leu Glu Lys Glu
820 825 830

AAC CGT GAA CTG GAA AAG ATC ATT GCT GAG AAA GAG GAG CGT GTC TCT 2544
Asn Arg Glu Leu Glu Lys Ile Ile Ala Glu Lys Glu Glu Arg Val Ser
835 840 845

GAA CTG CGC CAT CAA CTC CAG TCT CGG CAG CAG CTC CGC TCC CGG CGC 2592
Glu Leu Arg His Gln Leu Gln Ser Arg Gln Gln Leu Arg Ser Arg Arg
850 855 860

CAC CCA CCG ACA CCC CCA GAA CCC TCT GGG GGC CTG CCC AGG GGA CCC 2640
His Pro Pro Thr Pro Pro Glu Pro Ser Gly Gly Leu Pro Arg Gly Pro
865 870 875 880

CCT GAG CCC CCC GAC CGG CTT AGC TGT GAT GGG AGT CGA GTG CAT TTG 2688
Pro Glu Pro Pro Asp Arg Leu Ser Cys Asp Gly Ser Arg Val His Leu
885 890 895

CTT TAT AAG TGA 2700

Leu Tyr Lys *

(2) INFORMATION FOR SEQ ID NO: 71:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 899 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:

```
Met Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
 1               5                  10                  15

Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
            20                  25                  30

Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
        35                  40                  45

Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu
    50                  55                  60

Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys
65                  70                  75                  80

Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys
                85                  90                  95

Val Asn Arg Thr Pro His Ser Glu Arg Arg Ala Val Tyr Ile Gly Ala
            100                 105                 110

Leu Phe Pro Met Ser Gly Gly Trp Pro Gly Gly Gln Ala Cys Gln Pro
        115                 120                 125

Ala Val Glu Met Ala Leu Glu Asp Val Asn Ser Arg Arg Asp Ile Leu
    130                 135                 140

Pro Asp Tyr Glu Leu Lys Leu Ile His His Asp Ser Lys Cys Asp Pro
```

145 150 155 160

Gly Gln Ala Thr Lys Tyr Leu Tyr Glu Leu Leu Tyr Asn Asp Pro Ile
165 170 175

Lys Ile Ile Leu Met Pro Gly Cys Ser Ser Val Ser Thr Leu Val Ala
180 185 190

Glu Ala Ala Arg Met Trp Asn Leu Ile Val Leu Ser Tyr Gly Ser Ser
195 200 205

Ser Pro Ala Leu Ser Asn Arg Gln Arg Phe Pro Thr Phe Phe Arg Thr
210 215 220

His Pro Ser Ala Thr Leu His Asn Pro Thr Arg Val Lys Leu Phe Glu
225 230 235 240

Lys Trp Gly Trp Lys Lys Ile Ala Thr Ile Gln Gln Thr Thr Glu Val
245 250 255

Phe Thr Ser Thr Leu Asp Asp Leu Glu Glu Arg Val Lys Glu Ala Gly
260 265 270

Ile Glu Ile Thr Phe Arg Gln Ser Phe Phe Ser Asp Pro Ala Val Pro
275 280 285

Val Lys Asn Leu Lys Arg Gln Asp Ala Arg Ile Ile Val Gly Leu Phe
290 295 300

Tyr Glu Thr Glu Ala Arg Lys Val Phe Cys Glu Val Tyr Lys Glu Arg
305 310 315 320

Leu Phe Gly Lys Lys Tyr Val Trp Phe Leu Ile Gly Trp Tyr Ala Asp
325 330 335

Asn Trp Phe Lys Ile Tyr Asp Pro Ser Ile Asn Cys Thr Val Asp Glu
340 345 350

Met Thr Glu Ala Val Glu Gly His Ile Thr Thr Glu Ile Val Met Leu
355 360 365

Asn Pro Ala Asn Thr Arg Ser Ile Ser Asn Met Thr Ser Gln Glu Phe
370 375 380

```
Val Glu Lys Leu Thr Lys Arg Leu Lys Arg His Pro Glu Glu Thr Gly
385                 390                 395                 400

Gly Phe Gln Glu Ala Pro Leu Ala Tyr Asp Ala Ile Trp Ala Leu Ala
                405                 410                 415

Leu Ala Leu Asn Lys Thr Ser Gly Gly Gly Gly Arg Ser Gly Val Arg
            420                 425                 430

Leu Glu Asp Phe Asn Tyr Asn Asn Gln Thr Ile Thr Asp Gln Ile Tyr
        435                 440                 445

Arg Ala Met Asn Ser Ser Ser Phe Glu Gly Val Ser Gly His Val Val
    450                 455                 460

Phe Asp Ala Ser Gly Ser Arg Met Ala Trp Thr Leu Ile Glu Gln Leu
465                 470                 475                 480

Gln Gly Gly Ser Tyr Lys Lys Ile Gly Tyr Tyr Asp Ser Thr Lys Asp
                485                 490                 495

Asp Leu Ser Trp Ser Lys Thr Asp Lys Trp Ile Gly Gly Ser Pro Pro
            500                 505                 510

Ala Asp Gln Thr Leu Val Ile Lys Thr Phe Arg Phe Leu Ser Gln Lys
        515                 520                 525

Leu Phe Ile Ser Val Ser Val Leu Ser Ser Leu Gly Ile Val Leu Ala
    530                 535                 540

Val Val Cys Leu Ser Phe Asn Ile Tyr Asn Ser His Val Arg Tyr Ile
545                 550                 555                 560

Gln Asn Ser Gln Pro Asn Leu Asn Asn Leu Thr Ala Val Gly Cys Ser
                565                 570                 575

Leu Ala Leu Ala Ala Val Phe Pro Leu Gly Leu Asp Gly Tyr His Ile
            580                 585                 590

Gly Arg Asn Gln Phe Pro Phe Val Cys Gln Ala Arg Leu Trp Leu Leu
        595                 600                 605
```

Gly Leu Gly Phe Ser Leu Gly Tyr Gly Ser Met Phe Thr Lys Ile Trp
610 615 620

Trp Val His Thr Val Phe Thr Lys Lys Glu Glu Lys Lys Glu Trp Arg
625 630 635 640

Lys Thr Leu Glu Pro Trp Lys Leu Tyr Ala Thr Val Gly Leu Leu Val
645 650 655

Gly Met Asp Val Leu Thr Leu Ala Ile Trp Gln Ile Val Asp Pro Leu
660 665 670

His Arg Thr Ile Glu Thr Phe Ala Lys Glu Glu Pro Lys Glu Asp Ile
675 680 685

Asp Val Ser Ile Leu Pro Gln Leu Glu His Cys Ser Ser Arg Lys Met
690 695 700

Asn Thr Trp Leu Gly Ile Phe Tyr Gly Tyr Lys Gly Leu Leu Leu Leu
705 710 715 720

Leu Gly Ile Phe Leu Ala Tyr Glu Thr Lys Ser Val Ser Thr Glu Lys
725 730 735

Ile Asn Asp His Arg Ala Val Gly Met Ala Ile Tyr Asn Val Ala Val
740 745 750

Leu Cys Leu Ile Thr Ala Pro Val Thr Met Ile Leu Ser Ser Gln Gln
755 760 765

Asp Ala Ala Phe Ala Phe Ala Ser Leu Ala Ile Val Phe Ser Ser Tyr
770 775 780

Ile Thr Leu Val Val Leu Phe Val Pro Lys Met Arg Arg Leu Ile Thr
785 790 795 800

Arg Gly Glu Trp Gln Ser Glu Ala Gln Asp Thr Met Lys Thr Gly Ser
805 810 815

Ser Thr Asn Asn Asn Glu Glu Glu Lys Ser Arg Leu Leu Glu Lys Glu
820 825 830

Asn Arg Glu Leu Glu Lys Ile Ile Ala Glu Lys Glu Glu Arg Val Ser

835 840 845

Glu Leu Arg His Gln Leu Gln Ser Arg Gln Gln Leu Arg Ser Arg Arg
850 855 860

His Pro Pro Thr Pro Pro Glu Pro Ser Gly Gly Leu Pro Arg Gly Pro
865 870 875 880

Pro Glu Pro Pro Asp Arg Leu Ser Cys Asp Gly Ser Arg Val His Leu
885 890 895

Leu Tyr Lys

(2) INFORMATION FOR SEQ ID NO: 72:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2518 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..294

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:

ATG TTG CTG CTG CTG CTA CTG GCG CCA CTC TTC CTC CGC CCC CCG GGC 48
Met Leu Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
1 5 10 15

GCG GGC GGG GCG CAG ACC CCC AAC GCC ACC TCA GAA GGT TGC CAG ATC 96
Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile

20 25 30

ATA CAC CCG CCC TGG GAA GGG GGC ATC AGG TAC CGG GGC CTG ACT CGG 144
Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
35 40 45

GAC CAG GTG AAG GCT ATC AAC TTC CTG CCA GTG GAC TAT GAG ATT GAG 192
Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu
50 55 60

TAT GTG TGC CGG GGG GAG CGC GAG GTG GTG GGG CCC AAG GTC CGC AAG 240
Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys
65 70 75 80

TGC CTG GCC AAC GGC TCC TGG ACA GAT ATG GAC ACA CCC AGC CGC TGT 288
Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys
85 90 95

GTG TGA TCCAGGCCAA GCCACCAAGT ACCTATATGA GCTGCTCTAC AACGACCCTA 344
Val *

TCAAGATCAT CCTTATGCCT GGCTGCAGCT CTGTCTCCAC GCTGGTGGCT GAGGCTGCTA 404

GGATGTGGAA CCTCATTGTG CTTTCCTATG GCTCCAGCTC ACCAGCCCTG TCAAACCGGC 464

AGCGTTTCCC CACTTTCTTC CGAACGCACC CATCAGCCAC ACTCCACAAC CCTACCCGCG 524

TGAAACTCTT TGAAAAGTGG GGCTGGAAGA AGATTGCTAC CATCCAGCAG ACCACTGAGG 584

TCTTCACTTC GACTCTGGAC GACCTGGAGG AACGAGTGAA GGAGGCTGGA ATTGAGATTA 644

CTTTCCGCCA GAGTTTCTTC TCAGATCCAG CTGTGCCCGT CAAAAACCTG AAGCGCCAGG 704

ATGCCCGAAT CATCGTGGGA CTTTTCTATG AGACTGAAGC CCGGAAAGTT TTTTGTGAGG 764

TGTACAAGGA GCGTCTCTTT GGGAAGAAGT ACGTCTGGTT CCTCATTGGG TGGTATGCTG 824

ACAATTGGTT CAAGATCTAC GACCCTTCTA TCAACTGCAC AGTGGATGAG ATGACTGAGG 884

CGGTGGAGGG CCACATCACA ACTGAGATTG TCATGCTGAA TCCTGCCAAT ACCCGCAGCA 944

TTTCCAACAT GACATCCCAG GAATTTGTGG AGAAACTAAC CAAGCGACTG AAAAGACACC 1004

CTGAGGAGAC AGGAGGCTTC CAGGAGGCAC CGCTGGCCTA TGATGCCATC TGGGCCTTGG 1064

CACTGGCCCT GAACAAGACA TCTGGAGGAG GCGGCCGTTC TGGTGTGCGC CTGGAGGACT 1124

TCAACTACAA CAACCAGACC ATTACCGACC AAATCTACCG GGCAATGAAC TCTTCGTCCT 1184

TTGAGGGTGT CTCTGGCCAT GTGGTGTTTG ATGCCAGCGG CTCTCGGATG GCATGGACGC 1244

TTATCGAGCA GCTTCAGGGT GGCAGCTACA AGAAGATTGG CTACTATGAC AGCACCAAGG 1304

ATGATCTTTC CTGGTCCAAA ACAGATAAAT GGATTGGAGG GTCCCCCCCA GCTGACCAGA 1364

CCCTGGTCAT CAAGACATTC CGCTTCCTGT CACAGAAACT CTTTATCTCC GTCTCAGTTC 1424

TCTCCAGCCT GGGCATTGTC CTAGCTGTTG TCTGTCTGTC CTTTAACATC TACAACTCAC 1484

ATGTCCGTTA TATCCAGAAC TCACAGCCCA ACCTGAACAA CCTGACTGCT GTGGGCTGCT 1544

CACTGGCTTT AGCTGCTGTC TTCCCCCTGG GGCTCGATGG TTACCACATT GGGAGGAACC 1604

AGTTTCCTTT CGTCTGCCAG GCCCGCCTCT GGCTCCTGGG CCTGGGCTTT AGTCTGGGCT 1664

ACGGTTCCAT GTTCACCAAG ATTTGGTGGG TCCACACGGT CTTCACAAAG AAGGAAGAAA 1724

AGAAGGAGTG GAGGAAGACT CTGGAACCCT GGAAGCTGTA TGCCACAGTG GGCCTGCTGG 1784

TGGGCATGGA TGTCCTCACT CTCGCCATCT GGCAGATCGT GGACCCTCTG CACCGGACCA 1844

TTGAGACATT TGCCAAGGAG GAACCTAAGG AAGATATTGA CGTCTCTATT CTGCCCCAGC 1904

TGGAGCATTG CAGCTCCAGG AAGATGAATA CATGGCTTGG CATTTTCTAT GGTTACAAGG 1964

GGCTGCTGCT GCTGCTGGGA ATCTTCCTTG CTTATGAGAC CAAGAGTGTG TCCACTGAGA 2024

AGATCAATGA TCACCGGGCT GTGGGCATGG CTATCTACAA TGTGGCAGTC CTGTGCCTCA 2084

TCACTGCTCC TGTCACCATG ATTCTGTCCA GCCAGCAGGA TGCAGCCTTT GCCTTTGCCT 2144

CTCTTGCCAT AGTTTTCTCC TCCTATATCA CTCTTGTTGT GCTCTTTGTG CCCAAGATGC 2204

GCAGGCTGAT CACCCGAGGG GAATGGCAGT CGGAGGCGCA GGACACCATG AAGACAGGGT 2264

CATCGACCAA CAACAACGAG GAGGAGAAGT CCCGGCTGTT GGAGAAGGAG AACCGTGAAC 2324

TGGAAAAGAT CATTGCTGAG AAAGAGGAGC GTGTCTCTGA ACTGCGCCAT CAACTCCAGT 2384

CTCGGCAGCA GCTCCGCTCC CGGCGCCACC CACCGACACC CCCAGAACCC TCTGGGGGCC 2444

TGCCCAGGGG ACCCCCTGAG CCCCCCGACC GGCTTAGCTG TGATGGGAGT CGAGTGCATT 2504

TGCTTTATAA GTGA 2518

(2) INFORMATION FOR SEQ ID NO: 73:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 97 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 73:

Met Leu Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
1 5 10 15

Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
20 25 30

Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
35 40 45

Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu
50 55 60

Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys
65 70 75 80

Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys
85 90 95

Val

(2) INFORMATION FOR SEQ ID NO: 74:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2679 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..2679

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74:

```
ATG TTG CTG CTG CTG CTA CTG GCG CCA CTC TTC CTC CGC CCC CCG GGC        48
Met Leu Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
  1               5                  10                  15

GCG GGC GGG GCG CAG ACC CCC AAC GCC ACC TCA GAA GGT TGC CAG ATC        96
Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
            20                  25                  30

ATA CAC CCG CCC TGG GAA GGG GGC ATC AGG TAC CGG GGC CTG ACT CGG       144
Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
        35                  40                  45

GAC CAG GTG AAG GCT ATC AAC TTC CTG CCA GTG GAC TAT GAG ATT GAG       192
Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu
    50                  55                  60

TAT GTG TGC CGG GGG GAG CGC GAG GTG GTG GGG CCC AAG GTC CGC AAG       240
Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys
65                  70                  75                  80
```

```
TGC CTG GCC AAC GGC TCC TGG ACA GAT ATG GAC ACA CCC AGC CGC TGT      288
Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys
                85                  90                  95

GAA CGG CGC GCA GTG TAC ATC GGG GCA CTG TTT CCC ATG AGC GGG GGC      336
Glu Arg Arg Ala Val Tyr Ile Gly Ala Leu Phe Pro Met Ser Gly Gly
            100                 105                 110

TGG CCA GGG GGC CAG GCC TGC CAG CCC GCG GTG GAG ATG GCG CTG GAG      384
Trp Pro Gly Gly Gln Ala Cys Gln Pro Ala Val Glu Met Ala Leu Glu
        115                 120                 125

GAC GTG AAT AGC CGC AGG GAC ATC CTG CCG GAC TAT GAG CTC AAG CTC      432
Asp Val Asn Ser Arg Arg Asp Ile Leu Pro Asp Tyr Glu Leu Lys Leu
    130                 135                 140

ATC CAC CAC GAC AGC AAG TGT GAT CCA GGC CAA GCC ACC AAG TAC CTA      480
Ile His His Asp Ser Lys Cys Asp Pro Gly Gln Ala Thr Lys Tyr Leu
145                 150                 155                 160

TAT GAG CTG CTC TAC AAC GAC CCT ATC AAG ATC ATC CTT ATG CCT GGC      528
Tyr Glu Leu Leu Tyr Asn Asp Pro Ile Lys Ile Ile Leu Met Pro Gly
                165                 170                 175

TGC AGC TCT GTC TCC ACG CTG GTG GCT GAG GCT GCT AGG ATG TGG AAC      576
Cys Ser Ser Val Ser Thr Leu Val Ala Glu Ala Ala Arg Met Trp Asn
            180                 185                 190

CTC ATT GTG CTT TCC TAT GGC TCC AGC TCA CCA GCC CTG TCA AAC CGG      624
Leu Ile Val Leu Ser Tyr Gly Ser Ser Ser Pro Ala Leu Ser Asn Arg
        195                 200                 205

CAG CGT TTC CCC ACT TTC TTC CGA ACG CAC CCA TCA GCC ACA CTC CAC      672
Gln Arg Phe Pro Thr Phe Phe Arg Thr His Pro Ser Ala Thr Leu His
    210                 215                 220

AAC CCT ACC CGC GTG AAA CTC TTT GAA AAG TGG GGC TGG AAG AAG ATT      720
Asn Pro Thr Arg Val Lys Leu Phe Glu Lys Trp Gly Trp Lys Lys Ile
225                 230                 235                 240

GCT ACC ATC CAG CAG ACC ACT GAG GTC TTC ACT TCG ACT CTG GAC GAC      768
Ala Thr Ile Gln Gln Thr Thr Glu Val Phe Thr Ser Thr Leu Asp Asp
```

245 250 255

CTG GAG GAA CGA GTG AAG GAG GCT GGA ATT GAG ATT ACT TTC CGC CAG 816
Leu Glu Glu Arg Val Lys Glu Ala Gly Ile Glu Ile Thr Phe Arg Gln
260 265 270

AGT TTC TTC TCA GAT CCA GCT GTG CCC GTC AAA AAC CTG AAG CGC CAG 864
Ser Phe Phe Ser Asp Pro Ala Val Pro Val Lys Asn Leu Lys Arg Gln
275 280 285

GAT GCC CGA ATC ATC GTG GGA CTT TTC TAT GAG ACT GAA GCC CGG AAA 912
Asp Ala Arg Ile Ile Val Gly Leu Phe Tyr Glu Thr Glu Ala Arg Lys
290 295 300

GTT TTT TGT GAG GTG TAC AAG GAG CGT CTC TTT GGG AAG AAG TAC GTC 960
Val Phe Cys Glu Val Tyr Lys Glu Arg Leu Phe Gly Lys Lys Tyr Val
305 310 315 320

TGG TTC CTC ATT GGG TGG TAT GCT GAC AAT TGG TTC AAG ATC TAC GAC 1008
Trp Phe Leu Ile Gly Trp Tyr Ala Asp Asn Trp Phe Lys Ile Tyr Asp
325 330 335

CCT TCT ATC AAC TGC ACA GTG GAT GAG ATG ACT GAG GCG GTG GAG GGC 1056
Pro Ser Ile Asn Cys Thr Val Asp Glu Met Thr Glu Ala Val Glu Gly
340 345 350

CAC ATC ACA ACT GAG ATT GTC ATG CTG AAT CCT GCC AAT ACC CGC AGC 1104
His Ile Thr Thr Glu Ile Val Met Leu Asn Pro Ala Asn Thr Arg Ser
355 360 365

ATT TCC AAC ATG ACA TCC CAG GAA TTT GTG GAG AAA CTA ACC AAG CGA 1152
Ile Ser Asn Met Thr Ser Gln Glu Phe Val Glu Lys Leu Thr Lys Arg
370 375 380

CTG AAA AGA CAC CCT GAG GAG ACA GGA GGC TTC CAG GAG GCA CCG CTG 1200
Leu Lys Arg His Pro Glu Glu Thr Gly Gly Phe Gln Glu Ala Pro Leu
385 390 395 400

GCC TAT GAT GCC ATC TGG GCC TTG GCA CTG GCC CTG AAC AAG ACA TCT 1248
Ala Tyr Asp Ala Ile Trp Ala Leu Ala Leu Ala Leu Asn Lys Thr Ser
405 410 415

GGA GGA GGC GGC CGT TCT GGT GTG CGC CTG GAG GAC TTC AAC TAC AAC 1296

```
Gly Gly Gly Gly Arg Ser Gly Val Arg Leu Glu Asp Phe Asn Tyr Asn
            420                 425                 430

AAC CAG ACC ATT ACC GAC CAA ATC TAC CGG GCA ATG AAC TCT TCG TCC      1344
Asn Gln Thr Ile Thr Asp Gln Ile Tyr Arg Ala Met Asn Ser Ser Ser
        435                 440                 445

TTT GAG GGT GTC TCT GGC CAT GTG GTG TTT GAT GCC AGC GGC TCT CGG      1392
Phe Glu Gly Val Ser Gly His Val Val Phe Asp Ala Ser Gly Ser Arg
    450                 455                 460

ATG GCA TGG ACG CTT ATC GAG CAG CTT CAG GGT GGC AGC TAC AAG AAG      1440
Met Ala Trp Thr Leu Ile Glu Gln Leu Gln Gly Gly Ser Tyr Lys Lys
465                 470                 475                 480

ATT GGC TAC TAT GAC AGC ACC AAG GAT GAT CTT TCC TGG TCC AAA ACA      1488
Ile Gly Tyr Tyr Asp Ser Thr Lys Asp Asp Leu Ser Trp Ser Lys Thr
                485                 490                 495

GAT AAA TGG ATT GGA GGG TCC CCC CCA GCT GAC CAG ACC CTG GTC ATC      1536
Asp Lys Trp Ile Gly Gly Ser Pro Pro Ala Asp Gln Thr Leu Val Ile
            500                 505                 510

AAG ACA TTC CGC TTC CTG TCA CAG AAA CTC TTT ATC TCC GTC TCA GTT      1584
Lys Thr Phe Arg Phe Leu Ser Gln Lys Leu Phe Ile Ser Val Ser Val
        515                 520                 525

CTC TCC AGC CTG GGC ATT GTC CTA GCT GTT GTC TGT CTG TCC TTT AAC      1632
Leu Ser Ser Leu Gly Ile Val Leu Ala Val Val Cys Leu Ser Phe Asn
    530                 535                 540

ATC TAC AAC TCA CAT GTC CGT TAT ATC CAG AAC TCA CAG CCC AAC CTG      1680
Ile Tyr Asn Ser His Val Arg Tyr Ile Gln Asn Ser Gln Pro Asn Leu
545                 550                 555                 560

AAC AAC CTG ACT GCT GTG GGC TGC TCA CTG GCT TTA GCT GCT GTC TTC      1728
Asn Asn Leu Thr Ala Val Gly Cys Ser Leu Ala Leu Ala Ala Val Phe
                565                 570                 575

CCC CTG GGG CTC GAT GGT TAC CAC ATT GGG AGG AAC CAG TTT CCT TTC      1776
Pro Leu Gly Leu Asp Gly Tyr His Ile Gly Arg Asn Gln Phe Pro Phe
            580                 585                 590
```

```
GTC TGC CAG GCC CGC CTC TGG CTC CTG GGC CTG GGC TTT AGT CTG GGC      1824
Val Cys Gln Ala Arg Leu Trp Leu Leu Gly Leu Gly Phe Ser Leu Gly
        595                 600                 605

TAC GGT TCC ATG TTC ACC AAG ATT TGG TGG GTC CAC ACG GTC TTC ACA      1872
Tyr Gly Ser Met Phe Thr Lys Ile Trp Trp Val His Thr Val Phe Thr
    610                 615                 620

AAG AAG GAA GAA AAG AAG GAG TGG AGG AAG ACT CTG GAA CCC TGG AAG      1920
Lys Lys Glu Glu Lys Lys Glu Trp Arg Lys Thr Leu Glu Pro Trp Lys
625                 630                 635                 640

CTG TAT GCC ACA GTG GGC CTG CTG GTG GGC ATG GAT GTC CTC ACT CTC      1968
Leu Tyr Ala Thr Val Gly Leu Leu Val Gly Met Asp Val Leu Thr Leu
                645                 650                 655

GCC ATC TGG CAG ATC GTG GAC CCT CTG CAC CGG ACC ATT GAG ACA TTT      2016
Ala Ile Trp Gln Ile Val Asp Pro Leu His Arg Thr Ile Glu Thr Phe
            660                 665                 670

GCC AAG GAG GAA CCT AAG GAA GAT ATT GAC GTC TCT ATT CTG CCC CAG      2064
Ala Lys Glu Glu Pro Lys Glu Asp Ile Asp Val Ser Ile Leu Pro Gln
        675                 680                 685

CTG GAG CAT TGC AGC TCC AGG AAG ATG AAT ACA TGG CTT GGC ATT TTC      2112
Leu Glu His Cys Ser Ser Arg Lys Met Asn Thr Trp Leu Gly Ile Phe
    690                 695                 700

TAT GGT TAC AAG GGG CTG CTG CTG CTG CTG GGA ATC TTC CTT GCT TAT      2160
Tyr Gly Tyr Lys Gly Leu Leu Leu Leu Leu Gly Ile Phe Leu Ala Tyr
705                 710                 715                 720

GAG ACC AAG AGT GTG TCC ACT GAG AAG ATC AAT GAT CAC CGG GCT GTG      2208
Glu Thr Lys Ser Val Ser Thr Glu Lys Ile Asn Asp His Arg Ala Val
                725                 730                 735

GGC ATG GCT ATC TAC AAT GTG GCA GTC CTG TGC CTC ATC ACT GCT CCT      2256
Gly Met Ala Ile Tyr Asn Val Ala Val Leu Cys Leu Ile Thr Ala Pro
            740                 745                 750

GTC ACC ATG ATT CTG TCC AGC CAG CAG GAT GCA GCC TTT GCC TTT GCC      2304
Val Thr Met Ile Leu Ser Ser Gln Gln Asp Ala Ala Phe Ala Phe Ala
        755                 760                 765
```

```
TCT CTT GCC ATA GTT TTC TCC TCC TAT ATC ACT CTT GTT GTG CTC TTT      2352
Ser Leu Ala Ile Val Phe Ser Ser Tyr Ile Thr Leu Val Val Leu Phe
    770                 775                 780

GTG CCC AAG ATG CGC AGG CTG ATC ACC CGA GGG GAA TGG CAG TCG GAG      2400
Val Pro Lys Met Arg Arg Leu Ile Thr Arg Gly Glu Trp Gln Ser Glu
785                 790                 795                 800

GCG CAG GAC ACC ATG AAG ACA GGG TCA TCG ACC AAC AAC AAC GAG GAG      2448
Ala Gln Asp Thr Met Lys Thr Gly Ser Ser Thr Asn Asn Asn Glu Glu
                805                 810                 815

GAG AAG TCC CGG CTG TTG GAG AAG GAG AAC CGT GAA CTG GAA AAG ATC      2496
Glu Lys Ser Arg Leu Leu Glu Lys Glu Asn Arg Glu Leu Glu Lys Ile
            820                 825                 830

ATT GCT GAG AAA GAG GAG CGT GTC TCT GAA CTG CGC CAT CAA CTC CAG      2544
Ile Ala Glu Lys Glu Glu Arg Val Ser Glu Leu Arg His Gln Leu Gln
        835                 840                 845

TCT CGG CAG CAG CTC CGC TCC CGG CGC CAC CCA CCG ACA CCC CCA GAA      2592
Ser Arg Gln Gln Leu Arg Ser Arg Arg His Pro Pro Thr Pro Pro Glu
    850                 855                 860

CCC TCT GGG GGC CTG CCC AGG GGA CCC CCT GAG CCC CCC GAC CGG CTT      2640
Pro Ser Gly Gly Leu Pro Arg Gly Pro Pro Glu Pro Pro Asp Arg Leu
865                 870                 875                 880

AGC TGT GAT GGG AGT CGA GTG CAT TTG CTT TAT AAG TGA                  2679
Ser Cys Asp Gly Ser Arg Val His Leu Leu Tyr Lys  *
                885                 890
```

(2) INFORMATION FOR SEQ ID NO: 75:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 892 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:

```
Met Leu Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
 1               5                  10                  15

Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
            20                  25                  30

Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
        35                  40                  45

Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu
    50                  55                  60

Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys
65                  70                  75                  80

Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys
                85                  90                  95

Glu Arg Arg Ala Val Tyr Ile Gly Ala Leu Phe Pro Met Ser Gly Gly
            100                 105                 110

Trp Pro Gly Gly Gln Ala Cys Gln Pro Ala Val Glu Met Ala Leu Glu
        115                 120                 125

Asp Val Asn Ser Arg Arg Asp Ile Leu Pro Asp Tyr Glu Leu Lys Leu
    130                 135                 140

Ile His His Asp Ser Lys Cys Asp Pro Gly Gln Ala Thr Lys Tyr Leu
145                 150                 155                 160

Tyr Glu Leu Leu Tyr Asn Asp Pro Ile Lys Ile Ile Leu Met Pro Gly
                165                 170                 175

Cys Ser Ser Val Ser Thr Leu Val Ala Glu Ala Ala Arg Met Trp Asn
            180                 185                 190

Leu Ile Val Leu Ser Tyr Gly Ser Ser Ser Pro Ala Leu Ser Asn Arg
        195                 200                 205

Gln Arg Phe Pro Thr Phe Phe Arg Thr His Pro Ser Ala Thr Leu His
    210                 215                 220
```

```
Asn Pro Thr Arg Val Lys Leu Phe Glu Lys Trp Gly Trp Lys Lys Ile
225                 230                 235                 240

Ala Thr Ile Gln Gln Thr Thr Glu Val Phe Thr Ser Thr Leu Asp Asp
                245                 250                 255

Leu Glu Glu Arg Val Lys Glu Ala Gly Ile Glu Ile Thr Phe Arg Gln
            260                 265                 270

Ser Phe Phe Ser Asp Pro Ala Val Pro Val Lys Asn Leu Lys Arg Gln
        275                 280                 285

Asp Ala Arg Ile Ile Val Gly Leu Phe Tyr Glu Thr Glu Ala Arg Lys
    290                 295                 300

Val Phe Cys Glu Val Tyr Lys Glu Arg Leu Phe Gly Lys Lys Tyr Val
305                 310                 315                 320

Trp Phe Leu Ile Gly Trp Tyr Ala Asp Asn Trp Phe Lys Ile Tyr Asp
                325                 330                 335

Pro Ser Ile Asn Cys Thr Val Asp Glu Met Thr Glu Ala Val Glu Gly
            340                 345                 350

His Ile Thr Thr Glu Ile Val Met Leu Asn Pro Ala Asn Thr Arg Ser
        355                 360                 365

Ile Ser Asn Met Thr Ser Gln Glu Phe Val Glu Lys Leu Thr Lys Arg
    370                 375                 380

Leu Lys Arg His Pro Glu Glu Thr Gly Gly Phe Gln Glu Ala Pro Leu
385                 390                 395                 400

Ala Tyr Asp Ala Ile Trp Ala Leu Ala Leu Ala Leu Asn Lys Thr Ser
                405                 410                 415

Gly Gly Gly Gly Arg Ser Gly Val Arg Leu Glu Asp Phe Asn Tyr Asn
            420                 425                 430

Asn Gln Thr Ile Thr Asp Gln Ile Tyr Arg Ala Met Asn Ser Ser Ser
        435                 440                 445
```

Phe Glu Gly Val Ser Gly His Val Val Phe Asp Ala Ser Gly Ser Arg
450 455 460

Met Ala Trp Thr Leu Ile Glu Gln Leu Gln Gly Gly Ser Tyr Lys Lys
465 470 475 480

Ile Gly Tyr Tyr Asp Ser Thr Lys Asp Asp Leu Ser Trp Ser Lys Thr
485 490 495

Asp Lys Trp Ile Gly Gly Ser Pro Pro Ala Asp Gln Thr Leu Val Ile
500 505 510

Lys Thr Phe Arg Phe Leu Ser Gln Lys Leu Phe Ile Ser Val Ser Val
515 520 525

Leu Ser Ser Leu Gly Ile Val Leu Ala Val Val Cys Leu Ser Phe Asn
530 535 540

Ile Tyr Asn Ser His Val Arg Tyr Ile Gln Asn Ser Gln Pro Asn Leu
545 550 555 560

Asn Asn Leu Thr Ala Val Gly Cys Ser Leu Ala Leu Ala Ala Val Phe
565 570 575

Pro Leu Gly Leu Asp Gly Tyr His Ile Gly Arg Asn Gln Phe Pro Phe
580 585 590

Val Cys Gln Ala Arg Leu Trp Leu Leu Gly Leu Gly Phe Ser Leu Gly
595 600 605

Tyr Gly Ser Met Phe Thr Lys Ile Trp Trp Val His Thr Val Phe Thr
610 615 620

Lys Lys Glu Glu Lys Lys Glu Trp Arg Lys Thr Leu Glu Pro Trp Lys
625 630 635 640

Leu Tyr Ala Thr Val Gly Leu Leu Val Gly Met Asp Val Leu Thr Leu
645 650 655

Ala Ile Trp Gln Ile Val Asp Pro Leu His Arg Thr Ile Glu Thr Phe
660 665 670

Ala Lys Glu Glu Pro Lys Glu Asp Ile Asp Val Ser Ile Leu Pro Gln

675 680 685

Leu Glu His Cys Ser Ser Arg Lys Met Asn Thr Trp Leu Gly Ile Phe
690 695 700

Tyr Gly Tyr Lys Gly Leu Leu Leu Leu Leu Gly Ile Phe Leu Ala Tyr
705 710 715 720

Glu Thr Lys Ser Val Ser Thr Glu Lys Ile Asn Asp His Arg Ala Val
725 730 735

Gly Met Ala Ile Tyr Asn Val Ala Val Leu Cys Leu Ile Thr Ala Pro
740 745 750

Val Thr Met Ile Leu Ser Ser Gln Gln Asp Ala Ala Phe Ala Phe Ala
755 760 765

Ser Leu Ala Ile Val Phe Ser Ser Tyr Ile Thr Leu Val Val Leu Phe
770 775 780

Val Pro Lys Met Arg Arg Leu Ile Thr Arg Gly Glu Trp Gln Ser Glu
785 790 795 800

Ala Gln Asp Thr Met Lys Thr Gly Ser Ser Thr Asn Asn Asn Glu Glu
805 810 815

Glu Lys Ser Arg Leu Leu Glu Lys Glu Asn Arg Glu Leu Glu Lys Ile
820 825 830

Ile Ala Glu Lys Glu Glu Arg Val Ser Glu Leu Arg His Gln Leu Gln
835 840 845

Ser Arg Gln Gln Leu Arg Ser Arg Arg His Pro Pro Thr Pro Pro Glu
850 855 860

Pro Ser Gly Gly Leu Pro Arg Gly Pro Pro Glu Pro Pro Asp Arg Leu
865 870 875 880

Ser Cys Asp Gly Ser Arg Val His Leu Leu Tyr Lys
885 890

(2) INFORMATION FOR SEQ ID NO: 76:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2661 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..2661

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 76:

```
ATG TTG CTG CTG CTG CTA CTG GCG CCA CTC TTC CTC CGC CCC CCG GGC      48
Met Leu Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
 1               5                  10                  15

GCG GGC GGG GCG CAG ACC CCC AAC GCC ACC TCA GAA GGT TGC CAG ATC      96
Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
            20                  25                  30

ATA CAC CCG CCC TGG GAA GGG GGC ATC AGG TAC CGG GGC CTG ACT CGG     144
Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
        35                  40                  45

GAC CAG GTG AAG GCT ATC AAC TTC CTG CCA GTG GAC TAT GAG ATT GAG     192
Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu
    50                  55                  60

TAT GTG TGC CGG GGG GAG CGC GAG GTG GTG GGG CCC AAG GTC CGC AAG     240
Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys
65                  70                  75                  80

TGC CTG GCC AAC GGC TCC TGG ACA GAT ATG GAC ACA CCC AGC CGC TGT     288
```

```
Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys
                85                  90                  95

GTG AAT CGA ACG CCA CAC TCA GAA CGG CGC GCA GTG TAC ATC GGG GCA      336
Val Asn Arg Thr Pro His Ser Glu Arg Arg Ala Val Tyr Ile Gly Ala
            100                 105                 110

CTG TTT CCC GCG GTG GAG ATG GCG CTG GAG GAC GTG AAT AGC CGC AGG      384
Leu Phe Pro Ala Val Glu Met Ala Leu Glu Asp Val Asn Ser Arg Arg
        115                 120                 125

GAC ATC CTG CCG GAC TAT GAG CTC AAG CTC ATC CAC CAC GAC AGC AAG      432
Asp Ile Leu Pro Asp Tyr Glu Leu Lys Leu Ile His His Asp Ser Lys
    130                 135                 140

TGT GAT CCA GGC CAA GCC ACC AAG TAC CTA TAT GAG CTG CTC TAC AAC      480
Cys Asp Pro Gly Gln Ala Thr Lys Tyr Leu Tyr Glu Leu Leu Tyr Asn
145                 150                 155                 160

GAC CCT ATC AAG ATC ATC CTT ATG CCT GGC TGC AGC TCT GTC TCC ACG      528
Asp Pro Ile Lys Ile Ile Leu Met Pro Gly Cys Ser Ser Val Ser Thr
                165                 170                 175

CTG GTG GCT GAG GCT GCT AGG ATG TGG AAC CTC ATT GTG CTT TCC TAT      576
Leu Val Ala Glu Ala Ala Arg Met Trp Asn Leu Ile Val Leu Ser Tyr
            180                 185                 190

GGC TCC AGC TCA CCA GCC CTG TCA AAC CGG CAG CGT TTC CCC ACT TTC      624
Gly Ser Ser Ser Pro Ala Leu Ser Asn Arg Gln Arg Phe Pro Thr Phe
        195                 200                 205

TTC CGA ACG CAC CCA TCA GCC ACA CTC CAC AAC CCT ACC CGC GTG AAA      672
Phe Arg Thr His Pro Ser Ala Thr Leu His Asn Pro Thr Arg Val Lys
    210                 215                 220

CTC TTT GAA AAG TGG GGC TGG AAG AAG ATT GCT ACC ATC CAG CAG ACC      720
Leu Phe Glu Lys Trp Gly Trp Lys Lys Ile Ala Thr Ile Gln Gln Thr
225                 230                 235                 240

ACT GAG GTC TTC ACT TCG ACT CTG GAC GAC CTG GAG GAA CGA GTG AAG      768
Thr Glu Val Phe Thr Ser Thr Leu Asp Asp Leu Glu Glu Arg Val Lys
                245                 250                 255
```

```
GAG GCT GGA ATT GAG ATT ACT TTC CGC CAG AGT TTC TTC TCA GAT CCA      816
Glu Ala Gly Ile Glu Ile Thr Phe Arg Gln Ser Phe Phe Ser Asp Pro
            260                 265                 270

GCT GTG CCC GTC AAA AAC CTG AAG CGC CAG GAT GCC CGA ATC ATC GTG      864
Ala Val Pro Val Lys Asn Leu Lys Arg Gln Asp Ala Arg Ile Ile Val
        275                 280                 285

GGA CTT TTC TAT GAG ACT GAA GCC CGG AAA GTT TTT TGT GAG GTG TAC      912
Gly Leu Phe Tyr Glu Thr Glu Ala Arg Lys Val Phe Cys Glu Val Tyr
    290                 295                 300

AAG GAG CGT CTC TTT GGG AAG AAG TAC GTC TGG TTC CTC ATT GGG TGG      960
Lys Glu Arg Leu Phe Gly Lys Lys Tyr Val Trp Phe Leu Ile Gly Trp
305                 310                 315                 320

TAT GCT GAC AAT TGG TTC AAG ATC TAC GAC CCT TCT ATC AAC TGC ACA     1008
Tyr Ala Asp Asn Trp Phe Lys Ile Tyr Asp Pro Ser Ile Asn Cys Thr
                325                 330                 335

GTG GAT GAG ATG ACT GAG GCG GTG GAG GGC CAC ATC ACA ACT GAG ATT     1056
Val Asp Glu Met Thr Glu Ala Val Glu Gly His Ile Thr Thr Glu Ile
            340                 345                 350

GTC ATG CTG AAT CCT GCC AAT ACC CGC AGC ATT TCC AAC ATG ACA TCC     1104
Val Met Leu Asn Pro Ala Asn Thr Arg Ser Ile Ser Asn Met Thr Ser
        355                 360                 365

CAG GAA TTT GTG GAG AAA CTA ACC AAG CGA CTG AAA AGA CAC CCT GAG     1152
Gln Glu Phe Val Glu Lys Leu Thr Lys Arg Leu Lys Arg His Pro Glu
    370                 375                 380

GAG ACA GGA GGC TTC CAG GAG GCA CCG CTG GCC TAT GAT GCC ATC TGG     1200
Glu Thr Gly Gly Phe Gln Glu Ala Pro Leu Ala Tyr Asp Ala Ile Trp
385                 390                 395                 400

GCC TTG GCA CTG GCC CTG AAC AAG ACA TCT GGA GGA GGC GGC CGT TCT     1248
Ala Leu Ala Leu Ala Leu Asn Lys Thr Ser Gly Gly Gly Gly Arg Ser
                405                 410                 415

GGT GTG CGC CTG GAG GAC TTC AAC TAC AAC AAC CAG ACC ATT ACC GAC     1296
Gly Val Arg Leu Glu Asp Phe Asn Tyr Asn Asn Gln Thr Ile Thr Asp
            420                 425                 430
```

```
CAA ATC TAC CGG GCA ATG AAC TCT TCG TCC TTT GAG GGT GTC TCT GGC      1344
Gln Ile Tyr Arg Ala Met Asn Ser Ser Ser Phe Glu Gly Val Ser Gly
        435                 440                 445

CAT GTG GTG TTT GAT GCC AGC GGC TCT CGG ATG GCA TGG ACG CTT ATC      1392
His Val Val Phe Asp Ala Ser Gly Ser Arg Met Ala Trp Thr Leu Ile
    450                 455                 460

GAG CAG CTT CAG GGT GGC AGC TAC AAG AAG ATT GGC TAC TAT GAC AGC      1440
Glu Gln Leu Gln Gly Gly Ser Tyr Lys Lys Ile Gly Tyr Tyr Asp Ser
465                 470                 475                 480

ACC AAG GAT GAT CTT TCC TGG TCC AAA ACA GAT AAA TGG ATT GGA GGG      1488
Thr Lys Asp Asp Leu Ser Trp Ser Lys Thr Asp Lys Trp Ile Gly Gly
                485                 490                 495

TCC CCC CCA GCT GAC CAG ACC CTG GTC ATC AAG ACA TTC CGC TTC CTG      1536
Ser Pro Pro Ala Asp Gln Thr Leu Val Ile Lys Thr Phe Arg Phe Leu
            500                 505                 510

TCA CAG AAA CTC TTT ATC TCC GTC TCA GTT CTC TCC AGC CTG GGC ATT      1584
Ser Gln Lys Leu Phe Ile Ser Val Ser Val Leu Ser Ser Leu Gly Ile
        515                 520                 525

GTC CTA GCT GTT GTC TGT CTG TCC TTT AAC ATC TAC AAC TCA CAT GTC      1632
Val Leu Ala Val Val Cys Leu Ser Phe Asn Ile Tyr Asn Ser His Val
    530                 535                 540

CGT TAT ATC CAG AAC TCA CAG CCC AAC CTG AAC AAC CTG ACT GCT GTG      1680
Arg Tyr Ile Gln Asn Ser Gln Pro Asn Leu Asn Asn Leu Thr Ala Val
545                 550                 555                 560

GGC TGC TCA CTG GCT TTA GCT GCT GTC TTC CCC CTG GGG CTC GAT GGT      1728
Gly Cys Ser Leu Ala Leu Ala Ala Val Phe Pro Leu Gly Leu Asp Gly
                565                 570                 575

TAC CAC ATT GGG AGG AAC CAG TTT CCT TTC GTC TGC CAG GCC CGC CTC      1776
Tyr His Ile Gly Arg Asn Gln Phe Pro Phe Val Cys Gln Ala Arg Leu
            580                 585                 590

TGG CTC CTG GGC CTG GGC TTT AGT CTG GGC TAC GGT TCC ATG TTC ACC      1824
Trp Leu Leu Gly Leu Gly Phe Ser Leu Gly Tyr Gly Ser Met Phe Thr
```

595 600 605

AAG ATT TGG TGG GTC CAC ACG GTC TTC ACA AAG AAG GAA GAA AAG AAG 1872
Lys Ile Trp Trp Val His Thr Val Phe Thr Lys Lys Glu Glu Lys Lys
610 615 620

GAG TGG AGG AAG ACT CTG GAA CCC TGG AAG CTG TAT GCC ACA GTG GGC 1920
Glu Trp Arg Lys Thr Leu Glu Pro Trp Lys Leu Tyr Ala Thr Val Gly
625 630 635 640

CTG CTG GTG GGC ATG GAT GTC CTC ACT CTC GCC ATC TGG CAG ATC GTG 1968
Leu Leu Val Gly Met Asp Val Leu Thr Leu Ala Ile Trp Gln Ile Val
645 650 655

GAC CCT CTG CAC CGG ACC ATT GAG ACA TTT GCC AAG GAG GAA CCT AAG 2016
Asp Pro Leu His Arg Thr Ile Glu Thr Phe Ala Lys Glu Glu Pro Lys
660 665 670

GAA GAT ATT GAC GTC TCT ATT CTG CCC CAG CTG GAG CAT TGC AGC TCC 2064
Glu Asp Ile Asp Val Ser Ile Leu Pro Gln Leu Glu His Cys Ser Ser
675 680 685

AGG AAG ATG AAT ACA TGG CTT GGC ATT TTC TAT GGT TAC AAG GGG CTG 2112
Arg Lys Met Asn Thr Trp Leu Gly Ile Phe Tyr Gly Tyr Lys Gly Leu
690 695 700

CTG CTG CTG CTG GGA ATC TTC CTT GCT TAT GAG ACC AAG AGT GTG TCC 2160
Leu Leu Leu Leu Gly Ile Phe Leu Ala Tyr Glu Thr Lys Ser Val Ser
705 710 715 720

ACT GAG AAG ATC AAT GAT CAC CGG GCT GTG GGC ATG GCT ATC TAC AAT 2208
Thr Glu Lys Ile Asn Asp His Arg Ala Val Gly Met Ala Ile Tyr Asn
725 730 735

GTG GCA GTC CTG TGC CTC ATC ACT GCT CCT GTC ACC ATG ATT CTG TCC 2256
Val Ala Val Leu Cys Leu Ile Thr Ala Pro Val Thr Met Ile Leu Ser
740 745 750

AGC CAG CAG GAT GCA GCC TTT GCC TTT GCC TCT CTT GCC ATA GTT TTC 2304
Ser Gln Gln Asp Ala Ala Phe Ala Phe Ala Ser Leu Ala Ile Val Phe
755 760 765

TCC TCC TAT ATC ACT CTT GTT GTG CTC TTT GTG CCC AAG ATG CGC AGG 2352

```
Ser Ser Tyr Ile Thr Leu Val Val Leu Phe Val Pro Lys Met Arg Arg
    770                 775                 780

CTG ATC ACC CGA GGG GAA TGG CAG TCG GAG GCG CAG GAC ACC ATG AAG      2400
Leu Ile Thr Arg Gly Glu Trp Gln Ser Glu Ala Gln Asp Thr Met Lys
785                 790                 795                 800

ACA GGG TCA TCG ACC AAC AAC AAC GAG GAG GAG AAG TCC CGG CTG TTG      2448
Thr Gly Ser Ser Thr Asn Asn Asn Glu Glu Glu Lys Ser Arg Leu Leu
                805                 810                 815

GAG AAG GAG AAC CGT GAA CTG GAA AAG ATC ATT GCT GAG AAA GAG GAG      2496
Glu Lys Glu Asn Arg Glu Leu Glu Lys Ile Ile Ala Glu Lys Glu Glu
            820                 825                 830

CGT GTC TCT GAA CTG CGC CAT CAA CTC CAG TCT CGG CAG CAG CTC CGC      2544
Arg Val Ser Glu Leu Arg His Gln Leu Gln Ser Arg Gln Gln Leu Arg
        835                 840                 845

TCC CGG CGC CAC CCA CCG ACA CCC CCA GAA CCC TCT GGG GGC CTG CCC      2592
Ser Arg Arg His Pro Pro Thr Pro Pro Glu Pro Ser Gly Gly Leu Pro
    850                 855                 860

AGG GGA CCC CCT GAG CCC CCC GAC CGG CTT AGC TGT GAT GGG AGT CGA      2640
Arg Gly Pro Pro Glu Pro Pro Asp Arg Leu Ser Cys Asp Gly Ser Arg
865                 870                 875                 880

GTG CAT TTG CTT TAT AAG TGA                                         2661
Val His Leu Leu Tyr Lys  *
                885
```

(2) INFORMATION FOR SEQ ID NO: 77:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 886 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 77:

Met Leu Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
1 5 10 15

Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
20 25 30

Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
35 40 45

Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu
50 55 60

Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys
65 70 75 80

Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys
85 90 95

Val Asn Arg Thr Pro His Ser Glu Arg Arg Ala Val Tyr Ile Gly Ala
100 105 110

Leu Phe Pro Ala Val Glu Met Ala Leu Glu Asp Val Asn Ser Arg Arg
115 120 125

Asp Ile Leu Pro Asp Tyr Glu Leu Lys Leu Ile His His Asp Ser Lys
130 135 140

Cys Asp Pro Gly Gln Ala Thr Lys Tyr Leu Tyr Glu Leu Leu Tyr Asn
145 150 155 160

Asp Pro Ile Lys Ile Ile Leu Met Pro Gly Cys Ser Ser Val Ser Thr
165 170 175

Leu Val Ala Glu Ala Ala Arg Met Trp Asn Leu Ile Val Leu Ser Tyr
180 185 190

Gly Ser Ser Ser Pro Ala Leu Ser Asn Arg Gln Arg Phe Pro Thr Phe
195 200 205

Phe Arg Thr His Pro Ser Ala Thr Leu His Asn Pro Thr Arg Val Lys
210 215 220

Leu Phe Glu Lys Trp Gly Trp Lys Lys Ile Ala Thr Ile Gln Gln Thr

225 230 235 240

Thr Glu Val Phe Thr Ser Thr Leu Asp Asp Leu Glu Glu Arg Val Lys
245 250 255

Glu Ala Gly Ile Glu Ile Thr Phe Arg Gln Ser Phe Phe Ser Asp Pro
260 265 270

Ala Val Pro Val Lys Asn Leu Lys Arg Gln Asp Ala Arg Ile Ile Val
275 280 285

Gly Leu Phe Tyr Glu Thr Glu Ala Arg Lys Val Phe Cys Glu Val Tyr
290 295 300

Lys Glu Arg Leu Phe Gly Lys Lys Tyr Val Trp Phe Leu Ile Gly Trp
305 310 315 320

Tyr Ala Asp Asn Trp Phe Lys Ile Tyr Asp Pro Ser Ile Asn Cys Thr
325 330 335

Val Asp Glu Met Thr Glu Ala Val Glu Gly His Ile Thr Thr Glu Ile
340 345 350

Val Met Leu Asn Pro Ala Asn Thr Arg Ser Ile Ser Asn Met Thr Ser
355 360 365

Gln Glu Phe Val Glu Lys Leu Thr Lys Arg Leu Lys Arg His Pro Glu
370 375 380

Glu Thr Gly Gly Phe Gln Glu Ala Pro Leu Ala Tyr Asp Ala Ile Trp
385 390 395 400

Ala Leu Ala Leu Ala Leu Asn Lys Thr Ser Gly Gly Gly Gly Arg Ser
405 410 415

Gly Val Arg Leu Glu Asp Phe Asn Tyr Asn Asn Gln Thr Ile Thr Asp
420 425 430

Gln Ile Tyr Arg Ala Met Asn Ser Ser Ser Phe Glu Gly Val Ser Gly
435 440 445

His Val Val Phe Asp Ala Ser Gly Ser Arg Met Ala Trp Thr Leu Ile
450 455 460

```
Glu Gln Leu Gln Gly Gly Ser Tyr Lys Lys Ile Gly Tyr Tyr Asp Ser
465                 470                 475                 480

Thr Lys Asp Asp Leu Ser Trp Ser Lys Thr Asp Lys Trp Ile Gly Gly
                485                 490                 495

Ser Pro Pro Ala Asp Gln Thr Leu Val Ile Lys Thr Phe Arg Phe Leu
            500                 505                 510

Ser Gln Lys Leu Phe Ile Ser Val Ser Val Leu Ser Ser Leu Gly Ile
        515                 520                 525

Val Leu Ala Val Val Cys Leu Ser Phe Asn Ile Tyr Asn Ser His Val
    530                 535                 540

Arg Tyr Ile Gln Asn Ser Gln Pro Asn Leu Asn Asn Leu Thr Ala Val
545                 550                 555                 560

Gly Cys Ser Leu Ala Leu Ala Ala Val Phe Pro Leu Gly Leu Asp Gly
                565                 570                 575

Tyr His Ile Gly Arg Asn Gln Phe Pro Phe Val Cys Gln Ala Arg Leu
            580                 585                 590

Trp Leu Leu Gly Leu Gly Phe Ser Leu Gly Tyr Gly Ser Met Phe Thr
        595                 600                 605

Lys Ile Trp Trp Val His Thr Val Phe Thr Lys Lys Glu Glu Lys Lys
    610                 615                 620

Glu Trp Arg Lys Thr Leu Glu Pro Trp Lys Leu Tyr Ala Thr Val Gly
625                 630                 635                 640

Leu Leu Val Gly Met Asp Val Leu Thr Leu Ala Ile Trp Gln Ile Val
                645                 650                 655

Asp Pro Leu His Arg Thr Ile Glu Thr Phe Ala Lys Glu Glu Pro Lys
            660                 665                 670

Glu Asp Ile Asp Val Ser Ile Leu Pro Gln Leu Glu His Cys Ser Ser
        675                 680                 685
```

Arg Lys Met Asn Thr Trp Leu Gly Ile Phe Tyr Gly Tyr Lys Gly Leu
690 695 700

Leu Leu Leu Leu Gly Ile Phe Leu Ala Tyr Glu Thr Lys Ser Val Ser
705 710 715 720

Thr Glu Lys Ile Asn Asp His Arg Ala Val Gly Met Ala Ile Tyr Asn
725 730 735

Val Ala Val Leu Cys Leu Ile Thr Ala Pro Val Thr Met Ile Leu Ser
740 745 750

Ser Gln Gln Asp Ala Ala Phe Ala Phe Ala Ser Leu Ala Ile Val Phe
755 760 765

Ser Ser Tyr Ile Thr Leu Val Val Leu Phe Val Pro Lys Met Arg Arg
770 775 780

Leu Ile Thr Arg Gly Glu Trp Gln Ser Glu Ala Gln Asp Thr Met Lys
785 790 795 800

Thr Gly Ser Ser Thr Asn Asn Asn Glu Glu Glu Lys Ser Arg Leu Leu
805 810 815

Glu Lys Glu Asn Arg Glu Leu Glu Lys Ile Ile Ala Glu Lys Glu Glu
820 825 830

Arg Val Ser Glu Leu Arg His Gln Leu Gln Ser Arg Gln Gln Leu Arg
835 840 845

Ser Arg Arg His Pro Pro Thr Pro Pro Glu Pro Ser Gly Gly Leu Pro
850 855 860

Arg Gly Pro Pro Glu Pro Pro Asp Arg Leu Ser Cys Asp Gly Ser Arg
865 870 875 880

Val His Leu Leu Tyr Lys
885

(2) INFORMATION FOR SEQ ID NO: 78:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1692 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..1692

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 78:

```
ATG TTG CTG CTG CTG CTA CTG GCG CCA CTC TTC CTC CGC CCC CCG GGC      48
Met Leu Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
 1               5                  10                  15

GCG GGC GGG GCG CAG ACC CCC AAC GCC ACC TCA GAA GGT TGC CAG ATC      96
Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
            20                  25                  30

ATA CAC CCG CCC TGG GAA GGG GGC ATC AGG TAC CGG GGC CTG ACT CGG     144
Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
        35                  40                  45

GAC CAG GTG AAG GCT ATC AAC TTC CTG CCA GTG GAC TAT GAG ATT GAG     192
Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu
    50                  55                  60

TAT GTG TGC CGG GGG GAG CGC GAG GTG GTG GGG CCC AAG GTC CGC AAG     240
Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys
65                  70                  75                  80

TGC CTG GCC AAC GGC TCC TGG ACA GAT ATG GAC ACA CCC AGC CGC TGT     288
Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys
                85                  90                  95

GTC CGA ATC TGC TCC AAG TCT TAT TTG ACC ATT ACC GAC CAA ATC TAC     336
```

```
Val Arg Ile Cys Ser Lys Ser Tyr Leu Thr Ile Thr Asp Gln Ile Tyr
            100                 105                 110

CGG GCA ATG AAC TCT TCG TCC TTT GAG GGT GTC TCT GGC CAT GTG GTG      384
Arg Ala Met Asn Ser Ser Ser Phe Glu Gly Val Ser Gly His Val Val
        115                 120                 125

TTT GAT GCC AGC GGC TCT CGG ATG GCA TGG ACG CTT ATC GAG CAG CTT      432
Phe Asp Ala Ser Gly Ser Arg Met Ala Trp Thr Leu Ile Glu Gln Leu
    130                 135                 140

CAG GGT GGC AGC TAC AAG AAG ATT GGC TAC TAT GAC AGC ACC AAG GAT      480
Gln Gly Gly Ser Tyr Lys Lys Ile Gly Tyr Tyr Asp Ser Thr Lys Asp
145                 150                 155                 160

GAT CTT TCC TGG TCC AAA ACA GAT AAA TGG ATT GGA GGG TCC CCC CCA      528
Asp Leu Ser Trp Ser Lys Thr Asp Lys Trp Ile Gly Gly Ser Pro Pro
                165                 170                 175

GCT GAC CAG ACC CTG GTC ATC AAG ACA TTC CGC TTC CTG TCA CAG AAA      576
Ala Asp Gln Thr Leu Val Ile Lys Thr Phe Arg Phe Leu Ser Gln Lys
            180                 185                 190

CTC TTT ATC TCC GTC TCA GTT CTC TCC AGC CTG GGC ATT GTC CTA GCT      624
Leu Phe Ile Ser Val Ser Val Leu Ser Ser Leu Gly Ile Val Leu Ala
        195                 200                 205

GTT GTC TGT CTG TCC TTT AAC ATC TAC AAC TCA CAT GTC CGT TAT ATC      672
Val Val Cys Leu Ser Phe Asn Ile Tyr Asn Ser His Val Arg Tyr Ile
    210                 215                 220

CAG AAC TCA CAG CCC AAC CTG AAC AAC CTG ACT GCT GTG GGC TGC TCA      720
Gln Asn Ser Gln Pro Asn Leu Asn Asn Leu Thr Ala Val Gly Cys Ser
225                 230                 235                 240

CTG GCT TTA GCT GCT GTC TTC CCC CTG GGG CTC GAT GGT TAC CAC ATT      768
Leu Ala Leu Ala Ala Val Phe Pro Leu Gly Leu Asp Gly Tyr His Ile
                245                 250                 255

GGG AGG AAC CAG TTT CCT TTC GTC TGC CAG GCC CGC CTC TGG CTC CTG      816
Gly Arg Asn Gln Phe Pro Phe Val Cys Gln Ala Arg Leu Trp Leu Leu
            260                 265                 270
```

```
GGC CTG GGC TTT AGT CTG GGC TAC GGT TCC ATG TTC ACC AAG ATT TGG      864
Gly Leu Gly Phe Ser Leu Gly Tyr Gly Ser Met Phe Thr Lys Ile Trp
        275                 280                 285

TGG GTC CAC ACG GTC TTC ACA AAG AAG GAA GAA AAG AAG GAG TGG AGG      912
Trp Val His Thr Val Phe Thr Lys Lys Glu Glu Lys Lys Glu Trp Arg
    290                 295                 300

AAG ACT CTG GAA CCC TGG AAG CTG TAT GCC ACA GTG GGC CTG CTG GTG      960
Lys Thr Leu Glu Pro Trp Lys Leu Tyr Ala Thr Val Gly Leu Leu Val
305                 310                 315                 320

GGC ATG GAT GTC CTC ACT CTC GCC ATC TGG CAG ATC GTG GAC CCT CTG     1008
Gly Met Asp Val Leu Thr Leu Ala Ile Trp Gln Ile Val Asp Pro Leu
                325                 330                 335

CAC CGG ACC ATT GAG ACA TTT GCC AAG GAG GAA CCT AAG GAA GAT ATT     1056
His Arg Thr Ile Glu Thr Phe Ala Lys Glu Glu Pro Lys Glu Asp Ile
            340                 345                 350

GAC GTC TCT ATT CTG CCC CAG CTG GAG CAT TGC AGC TCC AGG AAG ATG     1104
Asp Val Ser Ile Leu Pro Gln Leu Glu His Cys Ser Ser Arg Lys Met
        355                 360                 365

AAT ACA TGG CTT GGC ATT TTC TAT GGT TAC AAG GGG CTG CTG CTG CTG     1152
Asn Thr Trp Leu Gly Ile Phe Tyr Gly Tyr Lys Gly Leu Leu Leu Leu
    370                 375                 380

CTG GGA ATC TTC CTT GCT TAT GAG ACC AAG AGT GTG TCC ACT GAG AAG     1200
Leu Gly Ile Phe Leu Ala Tyr Glu Thr Lys Ser Val Ser Thr Glu Lys
385                 390                 395                 400

ATC AAT GAT CAC CGG GCT GTG GGC ATG GCT ATC TAC AAT GTG GCA GTC     1248
Ile Asn Asp His Arg Ala Val Gly Met Ala Ile Tyr Asn Val Ala Val
                405                 410                 415

CTG TGC CTC ATC ACT GCT CCT GTC ACC ATG ATT CTG TCC AGC CAG CAG     1296
Leu Cys Leu Ile Thr Ala Pro Val Thr Met Ile Leu Ser Ser Gln Gln
            420                 425                 430

GAT GCA GCC TTT GCC TTT GCC TCT CTT GCC ATA GTT TTC TCC TCC TAT     1344
Asp Ala Ala Phe Ala Phe Ala Ser Leu Ala Ile Val Phe Ser Ser Tyr
        435                 440                 445
```

```
ATC ACT CTT GTT GTG CTC TTT GTG CCC AAG ATG CGC AGG CTG ATC ACC      1392
Ile Thr Leu Val Val Leu Phe Val Pro Lys Met Arg Arg Leu Ile Thr
    450                 455                 460

CGA GGG GAA TGG CAG TCG GAG GCG CAG GAC ACC ATG AAG ACA GGG TCA      1440
Arg Gly Glu Trp Gln Ser Glu Ala Gln Asp Thr Met Lys Thr Gly Ser
465                 470                 475                 480

TCG ACC AAC AAC AAC GAG GAG GAG AAG TCC CGG CTG TTG GAG AAG GAG      1488
Ser Thr Asn Asn Asn Glu Glu Glu Lys Ser Arg Leu Leu Glu Lys Glu
                485                 490                 495

AAC CGT GAA CTG GAA AAG ATC ATT GCT GAG AAA GAG GAG CGT GTC TCT      1536
Asn Arg Glu Leu Glu Lys Ile Ile Ala Glu Lys Glu Glu Arg Val Ser
            500                 505                 510

GAA CTG CGC CAT CAA CTC CAG TCT CGG CAG CAG CTC CGC TCC CGG CGC      1584
Glu Leu Arg His Gln Leu Gln Ser Arg Gln Gln Leu Arg Ser Arg Arg
        515                 520                 525

CAC CCA CCG ACA CCC CCA GAA CCC TCT GGG GGC CTG CCC AGG GGA CCC      1632
His Pro Pro Thr Pro Pro Glu Pro Ser Gly Gly Leu Pro Arg Gly Pro
    530                 535                 540

CCT GAG CCC CCC GAC CGG CTT AGC TGT GAT GGG AGT CGA GTG CAT TTG      1680
Pro Glu Pro Pro Asp Arg Leu Ser Cys Asp Gly Ser Arg Val His Leu
545                 550                 555                 560

CTT TAT AAG TGA                                                     1692
Leu Tyr Lys  *
```

(2) INFORMATION FOR SEQ ID NO: 79:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 563 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 79:

```
Met Leu Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
 1               5                  10                  15

Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
            20                  25                  30

Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
        35                  40                  45

Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu
    50                  55                  60

Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys
65                  70                  75                  80

Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys
                85                  90                  95

Val Arg Ile Cys Ser Lys Ser Tyr Leu Thr Ile Thr Asp Gln Ile Tyr
            100                 105                 110

Arg Ala Met Asn Ser Ser Ser Phe Glu Gly Val Ser Gly His Val Val
        115                 120                 125

Phe Asp Ala Ser Gly Ser Arg Met Ala Trp Thr Leu Ile Glu Gln Leu
    130                 135                 140

Gln Gly Gly Ser Tyr Lys Lys Ile Gly Tyr Tyr Asp Ser Thr Lys Asp
145                 150                 155                 160

Asp Leu Ser Trp Ser Lys Thr Asp Lys Trp Ile Gly Gly Ser Pro Pro
                165                 170                 175

Ala Asp Gln Thr Leu Val Ile Lys Thr Phe Arg Phe Leu Ser Gln Lys
            180                 185                 190

Leu Phe Ile Ser Val Ser Val Leu Ser Ser Leu Gly Ile Val Leu Ala
        195                 200                 205

Val Val Cys Leu Ser Phe Asn Ile Tyr Asn Ser His Val Arg Tyr Ile
    210                 215                 220
```

```
Gln Asn Ser Gln Pro Asn Leu Asn Asn Leu Thr Ala Val Gly Cys Ser
225                 230                 235                 240

Leu Ala Leu Ala Ala Val Phe Pro Leu Gly Leu Asp Gly Tyr His Ile
                245                 250                 255

Gly Arg Asn Gln Phe Pro Phe Val Cys Gln Ala Arg Leu Trp Leu Leu
            260                 265                 270

Gly Leu Gly Phe Ser Leu Gly Tyr Gly Ser Met Phe Thr Lys Ile Trp
        275                 280                 285

Trp Val His Thr Val Phe Thr Lys Lys Glu Glu Lys Lys Glu Trp Arg
    290                 295                 300

Lys Thr Leu Glu Pro Trp Lys Leu Tyr Ala Thr Val Gly Leu Leu Val
305                 310                 315                 320

Gly Met Asp Val Leu Thr Leu Ala Ile Trp Gln Ile Val Asp Pro Leu
                325                 330                 335

His Arg Thr Ile Glu Thr Phe Ala Lys Glu Glu Pro Lys Glu Asp Ile
            340                 345                 350

Asp Val Ser Ile Leu Pro Gln Leu Glu His Cys Ser Ser Arg Lys Met
        355                 360                 365

Asn Thr Trp Leu Gly Ile Phe Tyr Gly Tyr Lys Gly Leu Leu Leu Leu
    370                 375                 380

Leu Gly Ile Phe Leu Ala Tyr Glu Thr Lys Ser Val Ser Thr Glu Lys
385                 390                 395                 400

Ile Asn Asp His Arg Ala Val Gly Met Ala Ile Tyr Asn Val Ala Val
                405                 410                 415

Leu Cys Leu Ile Thr Ala Pro Val Thr Met Ile Leu Ser Ser Gln Gln
            420                 425                 430

Asp Ala Ala Phe Ala Phe Ala Ser Leu Ala Ile Val Phe Ser Ser Tyr
        435                 440                 445
```

```
Ile Thr Leu Val Val Leu Phe Val Pro Lys Met Arg Arg Leu Ile Thr
    450                 455                 460

Arg Gly Glu Trp Gln Ser Glu Ala Gln Asp Thr Met Lys Thr Gly Ser
465                 470                 475                 480

Ser Thr Asn Asn Asn Glu Glu Glu Lys Ser Arg Leu Leu Glu Lys Glu
                485                 490                 495

Asn Arg Glu Leu Glu Lys Ile Ile Ala Glu Lys Glu Glu Arg Val Ser
            500                 505                 510

Glu Leu Arg His Gln Leu Gln Ser Arg Gln Gln Leu Arg Ser Arg Arg
        515                 520                 525

His Pro Pro Thr Pro Pro Glu Pro Ser Gly Gly Leu Pro Arg Gly Pro
    530                 535                 540

Pro Glu Pro Pro Asp Arg Leu Ser Cys Asp Gly Ser Arg Val His Leu
545                 550                 555                 560

Leu Tyr Lys
```

(2) INFORMATION FOR SEQ ID NO: 80:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2602 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..318

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 80:

```
ATG TTG CTG CTG CTG CTA CTG GCG CCA CTC TTC CTC CGC CCC CCG GGC        48
Met Leu Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
 1               5                  10                  15

GCG GGC GGG GCG CAG ACC CCC AAC GCC ACC TCA GAA GGT TGC CAG ATC        96
Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
            20                  25                  30

ATA CAC CCG CCC TGG GAA GGG GGC ATC AGG TAC CGG GGC CTG ACT CGG       144
Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
        35                  40                  45

GAC CAG GTG AAG GCT ATC AAC TTC CTG CCA GTG GAC TAT GAG ATT GAT       192
Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Asp
    50                  55                  60

GAA TCG AAC GCC ACA CTC AGA ACG GCG CGC AGT GTA CAT CGG GGC ACT       240
Glu Ser Asn Ala Thr Leu Arg Thr Ala Arg Ser Val His Arg Gly Thr
65                  70                  75                  80

GTT TCC CAT GAG CGG GGG CTG GCC AGG GGG CCA GGC CTG CCA GCC CGC       288
Val Ser His Glu Arg Gly Leu Ala Arg Gly Pro Gly Leu Pro Ala Arg
                85                  90                  95

GGT GGA GAT GGC GCT GGA GGA CGT GAA TAG CCGCAGGGAC ATCCTGCCGG        338
Gly Gly Asp Gly Ala Gly Gly Arg Glu  *
            100                 105

ACTATGAGCT CAAGCTCATC CACCACGACA GCAAGTGTGA TCCAGGCCAA GCCACCAAGT    398

ACCTATATGA GCTGCTCTAC AACGACCCTA TCAAGATCAT CCTTATGCCT GGCTGCAGCT    458

CTGTCTCCAC GCTGGTGGCT GAGGCTGCTA GGATGTGGAA CCTCATTGTG CTTTCCTATG    518

GCTCCAGCTC ACCAGCCCTG TCAAACCGGC AGCGTTTCCC CACTTTCTTC CGAACGCACC    578

CATCAGCCAC ACTCCACAAC CCTACCCGCG TGAAACTCTT TGAAAAGTGG GGCTGGAAGA    638

AGATTGCTAC CATCCAGCAG ACCACTGAGG TCTTCACTTC GACTCTGGAC GACCTGGAGG    698
```

AACGAGTGAA GGAGGCTGGA ATTGAGATTA CTTTCCGCCA GAGTTTCTTC TCAGATCCAG 758

CTGTGCCCGT CAAAAACCTG AAGCGCCAGG ATGCCCGAAT CATCGTGGGA CTTTTCTATG 818

AGACTGAAGC CCGGAAAGTT TTTTGTGAGG TGTACAAGGA GCGTCTCTTT GGGAAGAAGT 878

ACGTCTGGTT CCTCATTGGG TGGTATGCTG ACAATTGGTT CAAGATCTAC GACCCTTCTA 938

TCAACTGCAC AGTGGATGAG ATGACTGAGG CGGTGGAGGG CCACATCACA ACTGAGATTG 998

TCATGCTGAA TCCTGCCAAT ACCCGCAGCA TTTCCAACAT GACATCCCAG GAATTTGTGG 1058

AGAAACTAAC CAAGCGACTG AAAAGACACC CTGAGGAGAC AGGAGGCTTC CAGGAGGCAC 1118

CGCTGGCCTA TGATGCCATC TGGGCCTTGG CACTGGCCCT GAACAAGACA TCTGGAGGAG 1178

GCGGCCGTTC TGGTGTGCGC CTGGAGGACT TCAACTACAA CAACCAGACC ATTACCGACC 1238

AAATCTACCG GGCAATGAAC TCTTCGTCCT TTGAGGGTGT CTCTGGCCAT GTGGTGTTTG 1298

ATGCCAGCGG CTCTCGGATG GCATGGACGC TTATCGAGCA GCTTCAGGGT GGCAGCTACA 1358

AGAAGATTGG CTACTATGAC AGCACCAAGG ATGATCTTTC CTGGTCCAAA ACAGATAAAT 1418

GGATTGGAGG GTCCCCCCCA GCTGACCAGA CCCTGGTCAT CAAGACATTC CGCTTCCTGT 1478

CACAGAAACT CTTTATCTCC GTCTCAGTTC TCTCCAGCCT GGGCATTGTC CTAGCTGTTG 1538

TCTGTCTGTC CTTTAACATC TACAACTCAC ATGTCCGTTA TATCCAGAAC TCACAGCCCA 1598

ACCTGAACAA CCTGACTGCT GTGGGCTGCT CACTGGCTTT AGCTGCTGTC TTCCCCCTGG 1658

GGCTCGATGG TTACCACATT GGGAGGAACC AGTTTCCTTT CGTCTGCCAG GCCCGCCTCT 1718

GGCTCCTGGG CCTGGGCTTT AGTCTGGGCT ACGGTTCCAT GTTCACCAAG ATTTGGTGGG 1778

TCCACACGGT CTTCACAAAG AAGGAAGAAA AGAAGGAGTG GAGGAAGACT CTGGAACCCT 1838

GGAAGCTGTA TGCCACAGTG GGCCTGCTGG TGGGCATGGA TGTCCTCACT CTCGCCATCT 1898

GGCAGATCGT GGACCCTCTG CACCGGACCA TTGAGACATT TGCCAAGGAG GAACCTAAGG 1958

AAGATATTGA CGTCTCTATT CTGCCCCAGC TGGAGCATTG CAGCTCCAGG AAGATGAATA 2018

CATGGCTTGG CATTTTCTAT GGTTACAAGG GGCTGCTGCT GCTGCTGGGA ATCTTCCTTG 2078

CTTATGAGAC CAAGAGTGTG TCCACTGAGA AGATCAATGA TCACCGGGCT GTGGGCATGG 2138

CTATCTACAA TGTGGCAGTC CTGTGCCTCA TCACTGCTCC TGTCACCATG ATTCTGTCCA 2198

GCCAGCAGGA TGCAGCCTTT GCCTTTGCCT CTCTTGCCAT AGTTTTCTCC TCCTATATCA 2258

CTCTTGTTGT GCTCTTTGTG CCCAAGATGC GCAGGCTGAT CACCCGAGGG GAATGGCAGT 2318

CGGAGGCGCA GGACACCATG AAGACAGGGT CATCGACCAA CAACAACGAG GAGGAGAAGT 2378

CCCGGCTGTT GGAGAAGGAG AACCGTGAAC TGGAAAAGAT CATTGCTGAG AAAGAGGAGC 2438

GTGTCTCTGA ACTGCGCCAT CAACTCCAGT CTCGGCAGCA GCTCCGCTCC CGGCGCCACC 2498

CACCGACACC CCCAGAACCC TCTGGGGGCC TGCCCAGGGG ACCCCCTGAG CCCCCCGACC 2558

GGCTTAGCTG TGATGGGAGT CGAGTGCATT TGCTTTATAA GTGA 2602

(2) INFORMATION FOR SEQ ID NO: 81:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 105 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 81:

Met Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
1 5 10 15

Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
20 25 30

Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
35 40 45

Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Asp

50 55 60

Glu Ser Asn Ala Thr Leu Arg Thr Ala Arg Ser Val His Arg Gly Thr
65 70 75 80

Val Ser His Glu Arg Gly Leu Ala Arg Gly Pro Gly Leu Pro Ala Arg
85 90 95

Gly Gly Asp Gly Ala Gly Gly Arg Glu
100 105

(2) INFORMATION FOR SEQ ID NO: 82:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2367 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..294

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 82:

ATG TTG CTG CTG CTG CTA CTG GCG CCA CTC TTC CTC CGC CCC CCG GGC 48
Met Leu Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
1 5 10 15

GCG GGC GGG GCG CAG ACC CCC AAC GCC ACC TCA GAA GGT TGC CAG ATC 96
Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
20 25 30

ATA CAC CCG CCC TGG GAA GGG GGC ATC AGG TAC CGG GGC CTG ACT CGG 144
Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg

35 40 45

GAC CAG GTG AAG GCT ATC AAC TTC CTG CCA GTG GAC TAT GAG ATT GAG 192
Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu
50 55 60

TAT GTG TGC CGG GGG GAG CGC GAG GTG GTG GGG CCC AAG GTC CGC AAG 240
Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys
65 70 75 80

TGC CTG GCC AAC GGC TCC TGG ACA GAT ATG GAC ACA CCC AGC CGC TGT 288
Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys
85 90 95

GTG TGA TCCAGGCCAA GCCACCAAGT ACCTATATGA GCTGCTCTAC AACGACCCTA 344
Val *

TCAAGATCAT CCTTATGCCT GGCTGCAGCT CTGTCTCCAC GCTGGTGGCT GAGGCTGCTA 404

GGATGTGGAA CCTCATTGTG CTTTCCTATG GCTCCAGCTC ACCAGCCCTG TCAAACCGGC 464

AGCGTTTCCC CACTTTCTTC CGAACGCACC CATCAGCCAC ACTCCACAAC CCTACCCGCG 524

TGAAACTCTT TGAAAAGTGG GGCTGGAAGA AGATTGCTAC CATCCAGCAG ACCACTGAGG 584

TCTTCACTTC GACTCTGGAC GACCTGGAGG AACGAGTGAA GGAGGCTGGA ATTGAGATTA 644

CTTTCCGCCA GAGTTTCTTC TCAGATCCAG CTGTGCCCGT CAAAAACCTG AAGCGCCAGG 704

ATGCCCGAAT CATCGTGGGA CTTTTCTATG AGACTGAAGC CCGGAAAGTT TTTTGTGAGG 764

TGTACAAGGA GCGTCTCTTT GGGAAGAAGT ACGTCTGGTT CCTCATTGGG TGGTATGCTG 824

ACAATTGGTT CAAGATCTAC GACCCTTCTA TCAACTGCAC AGTGGATGAG ATGACTGAGG 884

CGGTGGAGGG CCACATCACA ACTGAGATTG TCATGCTGAA TCCTGCCAAT ACCCGCAGCA 944

TTTCCAACAT GACATCCCAG GAATTTGTGG AGAAACTAAC CAAGCGACTG AAAAGACACC 1004

CTGAGGAGAC AGGAGGCTTC CAGGAGGCAC CGCTGGCCTA TGATGCCATC TGGGCCTTGG 1064

CACTGGCCCT GAACAAGACA TCTGGAGGAG GCGGCCGTTC TGGTGTGCGC CTGGAGGACT 1124

TCAACTACAA CAACCAGACC ATTACCGACC AAATCTACCG GGCAATGAAC TCTTCGTCCT 1184

TTGAGGGTGT CTCTGGCCAT GTGGTGTTTG ATGCCAGCGG CTCTCGGATG GCATGGACGC 1244

TTATCGAGCA GCTTCAGGGT GGCAGCTACA AGAAGATTGG CTACTATGAC AGCACCAAGG 1304

ATGATCTTTC CTGGTCCAAA ACAGATAAAT GGATTGTTAT ATCCAGAACT CACAGCCCAA 1364

CCTGAACAAC CTGACTGCTG TGGGCTGCTC ACTGGCTTTA GCTGCTGTCT TCCCCCTGGG 1424

GCTCGATGGT TACCACATTG GGAGGAACCA GTTTCCTTTC GTCTGCCAGG CCCGCCTCTG 1484

GCTCCTGGGC CTGGGCTTTA GTCTGGGCTA CGGTTCCATG TTCACCAAGA TTTGGTGGGT 1544

CCACACGGTC TTCACAAAGA AGGAAGAAAA GAAGGAGTGG AGGAAGACTC TGGAACCCTG 1604

GAAGCTGTAT GCCACAGTGG GCCTGCTGGT GGGCATGGAT GTCCTCACTC TCGCCATCTG 1664

GCAGATCGTG GACCCTCTGC ACCGGACCAT TGAGACATTT GCCAAGGAGG AACCTAAGGA 1724

AGATATTGAC GTCTCTATTC TGCCCCAGCT GGAGCATTGC AGCTCCAGGA AGATGAATAC 1784

ATGGCTTGGC ATTTTCTATG GTTACAAGGG GCTGCTGCTG CTGCTGGGAA TCTTCCTTGC 1844

TTATGAGACC AAGAGTGTGT CCACTGAGAA GATCAATGAT CACCGGGCTG TGGGCATGGC 1904

TATCTACAAT GTGGCAGTCC TGTGCCTCAT CACTGCTCCT GTCACCATGA TTCTGTCCAG 1964

CCAGCAGGAT GCAGCCTTTG CCTTTGCCTC TCTTGCCATA GTTTTCTCCT CCTATATCAC 2024

TCTTGTTGTG CTCTTTGTGC CCAAGATGCG CAGGCTGATC ACCCGAGGGG AATGGCAGTC 2084

GGAGGCGCAG GACACCATGA AGACAGGGTC ATCGACCAAC AACAACGAGG AGGAGAAGTC 2144

CCGGCTGTTG GAGAAGGAGA ACCGTGAACT GGAAAAGATC ATTGCTGAGA AAGAGGAGCG 2204

TGTCTCTGAA CTGCGCCATC AACTCCAGTC TCGGCAGCAG CTCCGCTCCC GGCGCCACCC 2264

ACCGACACCC CCAGAACCCT CTGGGGGCCT GCCCAGGGGA CCCCCTGAGC CCCCCGACCG 2324

GCTTAGCTGT GATGGGAGTC GAGTGCATTT GCTTTATAAG TGA 2367

(2) INFORMATION FOR SEQ ID NO: 83:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 97 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 83:

Met Leu Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
1 5 10 15

Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
20 25 30

Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
35 40 45

Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu
50 55 60

Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys
65 70 75 80

Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys
85 90 95

Val

(2) INFORMATION FOR SEQ ID NO: 84:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2489 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..1491

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 84:

```
ATG TTG CTG CTG CTG CTA CTG GCG CCA CTC TTC CTC CGC CCC CCG GGC      48
Met Leu Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
 1               5                  10                  15

GCG GGC GGG GCG CAG ACC CCC AAC GCC ACC TCA GAA GGT TGC CAG ATC      96
Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
            20                  25                  30

ATA CAC CCG CCC TGG GAA GGG GGC ATC AGG TAC CGG GGC CTG ACT CGG     144
Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
        35                  40                  45

GAC CAG GTG AAG GCT ATC AAC TTC CTG CCA GTG GAC TAT GAG ATT GAG     192
Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu
    50                  55                  60

TAT GTG TGC CGG GGG GAG CGC GAG GTG GTG GGG CCC AAG GTC CGC AAG     240
Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys
65                  70                  75                  80

TGC CTG GCC AAC GGC TCC TGG ACA GAT ATG GAC ACA CCC AGC CGC TGT     288
Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys
                85                  90                  95

GTC CGA ATC TGC TCC AAG TCT TAT TTG ACC CTG GAA AAT GGG AAG GTT     336
Val Arg Ile Cys Ser Lys Ser Tyr Leu Thr Leu Glu Asn Gly Lys Val
            100                 105                 110

TTC CTG ACG GGT GGG GAC CTC CCA GCT CTG GAC GGA GCC CGG GTG GAT     384
```

```
Phe Leu Thr Gly Gly Asp Leu Pro Ala Leu Asp Gly Ala Arg Val Asp
        115                 120                 125

TTC CGG TGT GAC CCC GAC TTC CAT CTG TGT GAT CCA GGC CAA GCC ACC      432
Phe Arg Cys Asp Pro Asp Phe His Leu Cys Asp Pro Gly Gln Ala Thr
    130                 135                 140

AAG TAC CTA TAT GAG CTG CTC TAC AAC GAC CCT ATC AAG ATC ATC CTT      480
Lys Tyr Leu Tyr Glu Leu Leu Tyr Asn Asp Pro Ile Lys Ile Ile Leu
145                 150                 155                 160

ATG CCT GGC TGC AGC TCT GTC TCC ACG CTG GTG GCT GAG GCT GCT AGG      528
Met Pro Gly Cys Ser Ser Val Ser Thr Leu Val Ala Glu Ala Ala Arg
                165                 170                 175

ATG TGG AAC CTC ATT GTG CTT TCC TAT GGC TCC AGC TCA CCA GCC CTG      576
Met Trp Asn Leu Ile Val Leu Ser Tyr Gly Ser Ser Ser Pro Ala Leu
            180                 185                 190

TCA AAC CGG CAG CGT TTC CCC ACT TTC TTC CGA ACG CAC CCA TCA GCC      624
Ser Asn Arg Gln Arg Phe Pro Thr Phe Phe Arg Thr His Pro Ser Ala
        195                 200                 205

ACA CTC CAC AAC CCT ACC CGC GTG AAA CTC TTT GAA AAG TGG GGC TGG      672
Thr Leu His Asn Pro Thr Arg Val Lys Leu Phe Glu Lys Trp Gly Trp
    210                 215                 220

AAG AAG ATT GCT ACC ATC CAG CAG ACC ACT GAG GTC TTC ACT TCG ACT      720
Lys Lys Ile Ala Thr Ile Gln Gln Thr Thr Glu Val Phe Thr Ser Thr
225                 230                 235                 240

CTG GAC GAC CTG GAG GAA CGA GTG AAG GAG GCT GGA ATT GAG ATT ACT      768
Leu Asp Asp Leu Glu Glu Arg Val Lys Glu Ala Gly Ile Glu Ile Thr
                245                 250                 255

TTC CGC CAG AGT TTC TTC TCA GAT CCA GCT GTG CCC GTC AAA AAC CTG      816
Phe Arg Gln Ser Phe Phe Ser Asp Pro Ala Val Pro Val Lys Asn Leu
            260                 265                 270

AAG CGC CAG GAT GCC CGA ATC ATC GTG GGA CTT TTC TAT GAG ACT GAA      864
Lys Arg Gln Asp Ala Arg Ile Ile Val Gly Leu Phe Tyr Glu Thr Glu
        275                 280                 285
```

```
GCC CGG AAA GTT TTT TGT GAG GTG TAC AAG GAG CGT CTC TTT GGG AAG      912
Ala Arg Lys Val Phe Cys Glu Val Tyr Lys Glu Arg Leu Phe Gly Lys
    290                 295                 300

AAG TAC GTC TGG TTC CTC ATT GGG TGG TAT GCT GAC AAT TGG TTC AAG      960
Lys Tyr Val Trp Phe Leu Ile Gly Trp Tyr Ala Asp Asn Trp Phe Lys
305                 310                 315                 320

ATC TAC GAC CCT TCT ATC AAC TGC ACA GTG GAT GAG ATG ACT GAG GCG     1008
Ile Tyr Asp Pro Ser Ile Asn Cys Thr Val Asp Glu Met Thr Glu Ala
                325                 330                 335

GTG GAG GGC CAC ATC ACA ACT GAG ATT GTC ATG CTG AAT CCT GCC AAT     1056
Val Glu Gly His Ile Thr Thr Glu Ile Val Met Leu Asn Pro Ala Asn
            340                 345                 350

ACC CGC AGC ATT TCC AAC ATG ACA TCC CAG GAA TTT GTG GAG AAA CTA     1104
Thr Arg Ser Ile Ser Asn Met Thr Ser Gln Glu Phe Val Glu Lys Leu
        355                 360                 365

ACC AAG CGA CTG AAA AGA CAC CCT GAG GAG ACA GGA GGC TTC CAG GAG     1152
Thr Lys Arg Leu Lys Arg His Pro Glu Glu Thr Gly Gly Phe Gln Glu
    370                 375                 380

GCA CCG CTG GCC TAT GAT GCC ATC TGG GCC TTG GCA CTG GCC CTG AAC     1200
Ala Pro Leu Ala Tyr Asp Ala Ile Trp Ala Leu Ala Leu Ala Leu Asn
385                 390                 395                 400

AAG ACA TCT GGA GGA GGC GGC CGT TCT GGT GTG CGC CTG GAG GAC TTC     1248
Lys Thr Ser Gly Gly Gly Gly Arg Ser Gly Val Arg Leu Glu Asp Phe
                405                 410                 415

AAC TAC AAC AAC CAG ACC ATT ACC GAC CAA ATC TAC CGG GCA ATG AAC     1296
Asn Tyr Asn Asn Gln Thr Ile Thr Asp Gln Ile Tyr Arg Ala Met Asn
            420                 425                 430

TCT TCG TCC TTT GAG GGT GTC TCT GGC CAT GTG GTG TTT GAT GCC AGC     1344
Ser Ser Ser Phe Glu Gly Val Ser Gly His Val Val Phe Asp Ala Ser
        435                 440                 445

GGC TCT CGG ATG GCA TGG ACG CTT ATC GAG CAG CTT CAG GGT GGC AGC     1392
Gly Ser Arg Met Ala Trp Thr Leu Ile Glu Gln Leu Gln Gly Gly Ser
    450                 455                 460
```

```
TAC AAG AAG ATT GGC TAC TAT GAC AGC ACC AAG GAT GAT CTT TCC TGG      1440
Tyr Lys Lys Ile Gly Tyr Tyr Asp Ser Thr Lys Asp Asp Leu Ser Trp
465                 470                 475                 480

TCC AAA ACA GAT AAA TGG ATT GTT ATA TCC AGA ACT CAC AGC CCA ACC      1488
Ser Lys Thr Asp Lys Trp Ile Val Ile Ser Arg Thr His Ser Pro Thr
                485                 490                 495

TGA ACAACCTGAC TGCTGTGGGC TGCTCACTGG CTTTAGCTGC TGTCTTCCCC           1541
*

CTGGGGCTCG ATGGTTACCA CATTGGGAGG AACCAGTTTC CTTTCGTCTG CCAGGCCCGC    1601

CTCTGGCTCC TGGGCCTGGG CTTTAGTCTG GGCTACGGTT CCATGTTCAC CAAGATTTGG    1661

TGGGTCCACA CGGTCTTCAC AAAGAAGGAA GAAAAGAAGG AGTGGAGGAA GACTCTGGAA    1721

CCCTGGAAGC TGTATGCCAC AGTGGGCCTG CTGGTGGGCA TGGATGTCCT CACTCTCGCC    1781

ATCTGGCAGA TCGTGGACCC TCTGCACCGG ACCATTGAGA CATTTGCCAA GGAGGAACCT    1841

AAGGAAGATA TTGACGTCTC TATTCTGCCC CAGCTGGAGC ATTGCAGCTC CAGGAAGATG    1901

AATACATGGC TTGGCATTTT CTATGGTTAC AAGGGGCTGC TGCTGCTGCT GGGAATCTTC    1961

CTTGCTTATG AGACCAAGAG TGTGTCCACT GAGAAGATCA ATGATCACCG GGCTGTGGGC    2021

ATGGCTATCT ACAATGTGGC AGTCCTGTGC CTCATCACTG CTCCTGTCAC CATGATTCTG    2081

TCCAGCCAGC AGGATGCAGC CTTTGCCTTT GCCTCTCTTG CCATAGTTTT CTCCTCCTAT    2141

ATCACTCTTG TTGTGCTCTT TGTGCCCAAG ATGCGCAGGC TGATCACCCG AGGGGAATGG    2201

CAGTCGGAGG CGCAGGACAC CATGAAGACA GGGTCATCGA CCAACAACAA CGAGGAGGAG    2261

AAGTCCCGGC TGTTGGAGAA GGAGAACCGT GAACTGGAAA AGATCATTGC TGAGAAAGAG    2321

GAGCGTGTCT CTGAACTGCG CCATCAACTC CAGTCTCGGC AGCAGCTCCG CTCCCGGCGC    2381

CACCCACCGA CACCCCCAGA ACCCTCTGGG GGCCTGCCCA GGGGACCCCC TGAGCCCCCC    2441
```

GACCGGCTTA GCTGTGATGG GAGTCGAGTG CATTTGCTTT ATAAGTGA 2489

(2) INFORMATION FOR SEQ ID NO: 85:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 496 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 85:

Met Leu Leu Leu Leu Leu Ala Pro Leu Phe Leu Arg Pro Pro Gly
1 5 10 15

Ala Gly Gly Ala Gln Thr Pro Asn Ala Thr Ser Glu Gly Cys Gln Ile
20 25 30

Ile His Pro Pro Trp Glu Gly Gly Ile Arg Tyr Arg Gly Leu Thr Arg
35 40 45

Asp Gln Val Lys Ala Ile Asn Phe Leu Pro Val Asp Tyr Glu Ile Glu
50 55 60

Tyr Val Cys Arg Gly Glu Arg Glu Val Val Gly Pro Lys Val Arg Lys
65 70 75 80

Cys Leu Ala Asn Gly Ser Trp Thr Asp Met Asp Thr Pro Ser Arg Cys
85 90 95

Val Arg Ile Cys Ser Lys Ser Tyr Leu Thr Leu Glu Asn Gly Lys Val
100 105 110

Phe Leu Thr Gly Gly Asp Leu Pro Ala Leu Asp Gly Ala Arg Val Asp
115 120 125

Phe Arg Cys Asp Pro Asp Phe His Leu Cys Asp Pro Gly Gln Ala Thr
130 135 140

Lys Tyr Leu Tyr Glu Leu Leu Tyr Asn Asp Pro Ile Lys Ile Ile Leu
145 150 155 160

Met Pro Gly Cys Ser Ser Val Ser Thr Leu Val Ala Glu Ala Ala Arg
165 170 175

Met Trp Asn Leu Ile Val Leu Ser Tyr Gly Ser Ser Ser Pro Ala Leu
180 185 190

Ser Asn Arg Gln Arg Phe Pro Thr Phe Phe Arg Thr His Pro Ser Ala
195 200 205

Thr Leu His Asn Pro Thr Arg Val Lys Leu Phe Glu Lys Trp Gly Trp
210 215 220

Lys Lys Ile Ala Thr Ile Gln Gln Thr Thr Glu Val Phe Thr Ser Thr
225 230 235 240

Leu Asp Asp Leu Glu Glu Arg Val Lys Glu Ala Gly Ile Glu Ile Thr
245 250 255

Phe Arg Gln Ser Phe Phe Ser Asp Pro Ala Val Pro Val Lys Asn Leu
260 265 270

Lys Arg Gln Asp Ala Arg Ile Ile Val Gly Leu Phe Tyr Glu Thr Glu
275 280 285

Ala Arg Lys Val Phe Cys Glu Val Tyr Lys Glu Arg Leu Phe Gly Lys
290 295 300

Lys Tyr Val Trp Phe Leu Ile Gly Trp Tyr Ala Asp Asn Trp Phe Lys
305 310 315 320

Ile Tyr Asp Pro Ser Ile Asn Cys Thr Val Asp Glu Met Thr Glu Ala
325 330 335

Val Glu Gly His Ile Thr Thr Glu Ile Val Met Leu Asn Pro Ala Asn
340 345 350

Thr Arg Ser Ile Ser Asn Met Thr Ser Gln Glu Phe Val Glu Lys Leu
355 360 365

Thr Lys Arg Leu Lys Arg His Pro Glu Glu Thr Gly Gly Phe Gln Glu
370 375 380

```
Ala Pro Leu Ala Tyr Asp Ala Ile Trp Ala Leu Ala Leu Ala Leu Asn
385                 390                 395                 400

Lys Thr Ser Gly Gly Gly Gly Arg Ser Gly Val Arg Leu Glu Asp Phe
                405                 410                 415

Asn Tyr Asn Asn Gln Thr Ile Thr Asp Gln Ile Tyr Arg Ala Met Asn
            420                 425                 430

Ser Ser Ser Phe Glu Gly Val Ser Gly His Val Val Phe Asp Ala Ser
        435                 440                 445

Gly Ser Arg Met Ala Trp Thr Leu Ile Glu Gln Leu Gln Gly Gly Ser
    450                 455                 460

Tyr Lys Lys Ile Gly Tyr Tyr Asp Ser Thr Lys Asp Asp Leu Ser Trp
465                 470                 475                 480

Ser Lys Thr Asp Lys Trp Ile Val Ile Ser Arg Thr His Ser Pro Thr
                485                 490                 495
```

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE 98/01947

A. CLASSIFICATION OF SUBJECT MATTER

IPC6: C07K 14/705
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC6: C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | EMBL Databas Genbank/DDBJ, accession no. Y11044, Grifa A. et al: "GABA (gamma-amino-butyric acid)neurotransmission: identification and fine mapping of human GABA-B receptor gene"; Biochem. Biophys. Res. Commun. 250:240-245(1998) -- | 1-6,11-13, 16-26 |
| X | Nature, Volume 386, March 1997, Klemens Kaupmann et al, "Expression cloning of GABAB receptors uncovers similarity to metabotropic glutamate receptors" page 239 - page 246 -- | 1-6,11-13, 16-26 |
| X | WO 9746675 A1 (NOVARTIS AG), 11 December 1997 (11.12.97), page 50 - page 56, claim 4 -- | 1-6,11-13, 16-26 |

[X] Further documents are listed in the continuation of Box C. [X] See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" erlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 23 March 1999 | 26 -03- 1999 |
| Name and mailing address of the ISA<br>Swedish Patent Office<br>Box 5055, S-102 42 STOCKHOLM<br>Facsimile No. +46 8 666 02 86 | Authorized officer<br>Patrick Andersson<br>Telephone No. +46 8 782 25 00 |

Form PCT ISA 210 (second sheet) (July 1992)

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE 98/01947

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| P,X | Dialog Information Services, File 34, SciSearch, Dialog accession no. 07157158, Goei VL et al: "Human gamma-aminobutyric acid B receptor gene: Complementary DNA cloning, expression, chromosomal location, and genomic organization", Biological Psychiatry, 1998, V44, N8(OCT 15), p659-666 -- | 1-6,11-13, 16-26 |
| A | Brain Research Bulletin, Volume 38, No 6, 1995, Robert J. Washabau et al, "GABA Receptros in the Dorsal Motor Nucleus of the Vagus Influence Feline Lower Esophageal Sphincter and Gastric Function" page 587 - page 594 -- -------- | 1-6,11-13, 16-26 |

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE 98/01947

**Box I** Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II** Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See next page

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

According to PCT rule 13.2, an international application shall relate to one invention only or a group of inventions linked by one or more of the same or corresponding "special technical features", i.e. features that define a contribution which each of the inventions makes over the prior art.

The claimed invention relates to a nucleic acid sequence encoding a GABAB-receptor from either human or canine origin. A possible special technical features could have been nucleotide sequence encoding a GABAB-receptor. However a nucleotide sequence encoding a human GABAB-receptor is known from Grifa A. et. al., see the search report. Thus, the present application is considered to contain the following independent inventions:

Invention 1, claims 3-4, and the parts of claims 1-2 and 16-26 relating to claims 3-4:A human GABAB-receptor 1a, and related items,

Invention 2, claims 5-6, and the parts of claims 1-2 and 16-26 relating to claims 5-6:A human GABAB-receptor 1b, and related items,

Invention 3, claims 7-8, and the parts of claims 1-2 and 16-26 relating to claims 7-8:A human GABAB-receptor 1c, and related items,

Invention 4, claims 9-10, and the parts of claims 1-2 and 16-26 relating to claims 9-10:A human GABAB-receptor 1d, and related items,

Invention 5, claims 12-13, and the parts of claims 1, 11 and 16-26 relating to claims 12-13:A canine GABAB-receptor 1a, and related items.

Invention 6, claims 14-15, and the parts of claims 1, 11 and 16-26 relating to claims 14-15:A canine GABAB-receptor 1c, and related items.

The search has been limited to inventions 1, 2 and 5.

# INTERNATIONAL SEARCH REPORT

Information on patent family members

02/03/99

International application No.
PCT/SE 98/01947

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| WO 9746675 A1 | 11/12/97 | AU 2028497 A | 05/01/98 |

Form PCT/ISA/210 (patent family annex) (July 1992)